Multi-omics Methods to unravel Microbial Diversity in Fermentation of Riesling Wines

Vom Fachbereich Chemie der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades “Doktor der Naturwissenschaften” genehmigte Dissertation

vorgelegt von

MSc. Kimmo Sirén

Angefertigt im Arbeitskreis von

Prof. Dr. Ulrich Fischer

Datum der wissenschaftlichen Aussprache:	08.03.2019
Vorsitzender:	Prof. Dr. Werner Thiel
Berichterstatter:	Prof. Dr. Elke Richling
Berichterstatter:	Prof. Dr. Ulrich Fischer

Kaiserslautern, 2019 - D386

This work was funded by Marie Skłodowska-Curie Innovative Training Network, “MicroWine”, Horizon 2020 Programme of the European Commission. This work was performed between August 2015 and February 2019 in the group of Prof. Dr. Ulrich Fischer in Institute for Viticulture and Oenology in DLR Rheinpfalz, Neustadt an der Weinstraße, Germany.

Promotionskommission:

Vorsitzender:	Prof. Dr. Werner Thiel
Berichterstatter:	Prof. Dr. Elke Richling
Berichterstatter:	Prof. Dr. Ulrich Fischer
Eingereicht am:	01.02.2019
Datum der wissenschaftlichen Aussprache:	08.03.2019

Table of contents

Declaration of Originality

I hereby declare that this thesis represents my own work, it was performed without help of third parties and no other sources of information have been used than those indicated. All informations that were gathered from released or unreleased sources were marked definitely as such.

In Kaiserslautern, _____

Acknowledgements

This work has been carried out at various locations but mainly at the Institute for Viticulture and Oenology in DLR Rheinpfalz and was made possible by the amazing funding of “MicroWine” the Marie Skłodowska-Curie Innovative Training Network, “MicroWine”, Horizon 2020, Programme of the European Commission.

I would like to thank my supervisor Prof. Dr. Ulrich Fischer for generous support throughout my studies. I was very privileged to have a lot of freedom to work on my own ideas but was always supported if I started to bang my head against the wall. Thank you for trusting me to be able to carry out such a big projects.

Thank you Prof. Dr. Elke Richling for kindly accepting to review my thesis for being in the thesis defence committee.

Thank you Prof. Dr. Werner Thiel for kindly accepting the position of the committee chair.

Thank you Prof. Dr. Tom Gilbert for initiating this marvellous network but also for giving me a workplace and inviting me to the Christmas parties and tremendous support throughout the project.

I would like to thank all the people that I met and collaborated with through various collaborations.

Thank you Prof. Dr. Vicente Ferreira, Inês and Eduardo and the whole Ferreira lab for having me around in Zaragoza. Especially thank you Vicente for having truly inspired me during that great night out. Thank you Andreas Dunkel and Prof. Dr. Thomas Hoffmann for hosting me in Freising. Thank you Hentie and Sofie for great help and discussions. Thank you Markus Herderich, Luca Nicolotti for hosting me at AWRI and allowing me to use your setup. Big thank you Wes and Alex for making me feel home and see really cool wine projects. Cheers to Lauren and Melanie for the cool time and trips in Australia.

I would also like to thank all other co-workers at the DLR Rheinpfalz for amazing moments and support, especially Jochen, Sebastian, Matthias, Engela, Michael W, Michael Z, Patricia, Daniel, Max and everyone else helping me with daily work life and projects. Big thank you especially to Jochen for listening and (not listening) my rants and

helping me with all this German. Thank you Sebastian and Matthias for the memorable lunch moments. Thank you Judith and Petra for helping me with all the admin issues. Thank you Pascal having great ideas and wine discussions. Thank you Hannah for being harvest help. Thank you all the people involved in the picking of grapes. Thank you all the people involved in the sensory analysis of the wines. Thank you Sandra for taking great care of the sensory panel. Thank you Anni for being my only student helper for the harvest.

I would like to thank all the people I in the EvoGenomics section, Anna, Lara, Inger, Inge, Marcela, Shyam, Miyako, Mikkel, Felipe. Special thank you Christian for being amazing host in KU and all the good moments and wine. Thank you Sarah for being amazing host in KU and all the good moments and wine.

I would like to thank Lars and Lea, for managing such a great Microwine project. Further thanks to all Microwine fellows and their PIs: Alex, Chrats, Franzi, Gio, Inês, Ingrid, Ifigeneia, Lei, Marc, Moni, Prashy, Rui, Sarah and Taneli for amazing times and travels. Especially I would like to thank the people involved around my Riesling projects – Sarah, Chrats, Inês and Franzi. I really have enjoyed collaborating with you.

Thank you Sarah for doing all the boring but ever so important parts in the collaborations and other projects, and especially listening to my rants.

Also I cannot thank enough everyone involved in my countless Riesling projects and all my co-authors in the publications.

All of this work could have not been possible without the contribution of the 10 wine estates donating huge amount of grapes. You are amazing.

Thank you Kai, Markus, Madeline, Johannes, Nicola, Sophie, Wes, Alex and countless others non-academics for keeping the wine a central theme. Thank you Sascha and Christian and the whole team at Meininger Verlag for inviting me to such a many interesting tastings. Thank you all the wine drinkers and wine enthusiasts and my friends throughout the world and wine world.

Thank you my parents and my brother and his family for their love and support.

Thank you Tiia for continuously supporting therefore making this possible.

Summary

Wine and alcoholic fermentations are complex and fascinating ecosystems. Wine aroma is shaped by the wine's chemical compositions, in which both microbes and grape constituents play crucial roles. Activities of the microbial community impact the sensory properties of the final product, therefore, the characterisation of microbial diversity is essential in understanding and predicting sensory properties of wine. Characterisation has been challenging with traditional approaches, where microbes are isolated and therefore analyzed outside from their natural environment. This causes a bias in the observed microbial composition structure. In addition, true community interactions cannot be studied using isolates. Furthermore, the multiplex ties between wine chemical and sensory compositions remain evasive due to their multivariate and nonlinear nature. Therefore, the sensorial outcome arising from different microbial communities has remained inconclusive.

In this thesis, microbial diversity during Riesling wine fermentations is investigated with the aim to understand the roles of microbial communities during fermentations and their links to sensory properties. With the advancement of high-throughput tools based 'omic methods, such as next-generation sequencing (NGS) technologies, it is now possible to study microbial communities and their functions without isolation by culturing. This developing field and its potential to wine community is reviewed in Chapter 1. The standardisation of methods remains challenging in the field. DNA extraction is a key step in capturing the microbial diversity in samples for generating NGS data, therefore, DNA extraction methods are evaluated in Chapter 2. In Chapter 3, machine learning is utilized in guiding raw data mining generated by the untargeted GC-MS analysis. This step is crucial in order to take full advantages of the large scope of data generated by 'omic methods. These lay a solid foundation for Chapters 4 and 5 where microbial community structures and their outputs - chemical and sensory compositions are studied by using approaches and tools based on multiple 'omics methods.

The results of this thesis show first that by using novel statistical approaches, it is possible to extract meaningful information from heterogeneous biological, chemical and sensorial data. Secondly, results suggest that the variation in wine aroma, might be related to microbial interactions taking place not only inside a single community, but also the

interactions between communities, such as vineyard and winery communities. Therefore, the true sensory expression of terroir might be masked by the interaction between two microbial communities, although more work is needed to uncover this potential relationship. Such potential interaction mechanisms were uncovered between non-*Saccharomyces* yeast and bacteria in this work and unexpected novel bacterial growth was observed during alcohol fermentation. This suggests new layers in understanding of wine fermentations. In the future, multi-omic approaches could be applied to identify biological pathways leading to specific wine aroma as well as investigate the effects upon specific winemaking conditions. These results are relevant not just for the wine industry, but also to other industries where complex microbial networks are important. As such, the approaches presented in this thesis might find widely use in the food industry.

Zusammenfassung

Wein im Allgemeinen und die alkoholische Gärung im Speziellen basieren in ihrem Entstehen auf komplexen und faszinierenden Ökosystemen. Das Weinaroma wird maßgeblich durch die chemische Zusammensetzung des Weines bestimmt, welche hauptsächlich von Verbindungen aus der Traube und solchen mikrobiologischen Ursprungs geprägt ist. Die mikrobielle Aktivität hat einen sowohl positiven als auch potenziell negativen sensorischen Einfluss auf das fermentierte Endprodukt, wodurch die Charakterisierung der mikrobiellen Diversität essentiell für das Verständnis und ein mögliche Vorhersage der sensorischen Eigenschaften von Weinen ist. Diese Charakterisierung ist mit traditionellen Methoden, bei welchen Mikroorganismen isoliert und somit außerhalb ihrer natürlichen Umgebung kultiviert und analysiert werden, schwierig. Diese Methoden sind mit einem systematischen Fehler hinsichtlich der der Zusammensetzung der Mikrobenpopulation behaftet, da es zahlreiche relevante Mikroorganismen gibt, die zwar lebensfähig, nicht aber kultivierbar sind. Des Weiteren sind die komplexen Zusammenhänge zwischen den sensorischen Eigenschaften und der chemischen Zusammensetzung des Weins, aufgrund ihrer multivariaten und nicht-linearen Natur, nur unzulänglich erfassbar. Der Einfluss ganz bestimmter mikrobiologischer Populationen auf die sensorischen Eigenschaften ist daher bisher nur ansatzweise und unzureichend aufgeklärt

In der vorliegenden Doktorarbeit wird die mikrobielle Diversität während der Gärung untersucht um die Rolle der Mikrobenpopulation und deren Zusammenhang mit sensorischen Eigenschaften des Weins zu beleuchten. Die Fortschritte in der Entwicklung von Omics-Methoden, wie etwa die next-generation sequencing (NGS) Technologie, ermöglichen einen hohen Probendurchsatz. Dies erlaubt heutzutage Mikrobenpopulationen und deren Funktionen ohne vorherige Isolierung und Kultivierung direkt zu untersuchen. Diese neuartigen Forschungsansätze und ihr Potential für die Weinforschung werden in Kapitel 1 vorgestellt. Dabei stellt die zwingend notwendige Standardisierung der Methoden eine große Herausforderung dar. Da die DNA-Extraktion ein Schlüsselschritt bei der Erfassung der mikrobiellen Diversität für die NGS-Datenerhebung ist, wird in Kapitel 2 ein Überblick über DNA-Extraktionsmethoden gegeben und sie in ihrer Leistungsfähigkeit verglichen. Kapitel 3 stellt eine

neuentwickelte Anwendung von maschinellen Lernen zur Durchführung von Data-Mining von Rohdaten aus nicht zielgerichteten GC-MS Analysen vor. Dieser Schritt ist im Hinblick auf die riesigen Datenmengen, die bei Omics-Methoden anfallen, unerlässlich. Hier wird auch das Fundament für Kapitel 4 und 5 gelegt, in welchen die strukturelle Zusammensetzung und deren Output, die chemische Zusammensetzung und sensorische Ausprägung, mittels Methoden und Werkzeugen, die auf multiplen Omics-Methoden beruhen, untersucht werden.

Die Ergebnisse der vorliegenden Arbeit zeigen die Möglichkeiten auf, mittels neuen statistischen Ansätzen sinnvolle Information von heterogenen biologischen, chemischen und sensorischen Daten zu extrahieren. Des Weiteren deuten die Ergebnisse darauf hin, dass Unterschiede im Weinaroma nicht nur auf Interaktionen innerhalb einer einzelnen Populationen von Mikroorganismen zurückzuführen sind, sondern auch auf solche zwischen Populationen, die sowohl aus dem Weinberg selbst als auch dem Weinkeller stammen. Der wahre sensorische Ausdruck eines Terroirs, sprich der geographischen Herkunft eines Weines, könnte durch die Interaktionen zwischen mikrobiologischen Populationen maskiert oder sogar komplementiert werden. Sicherlich sind weitere Studien notwendig sind, um diese Hypothese zu bestätigen und potentielle Beziehung aufzuklären. In dieser Arbeit wurden potentielle Interaktionsmechanismen zwischen Nicht-*Saccharomyces* und Bakterien aufgedeckt und unerwartetes Bakterienwachstum während der alkoholischen Gärung führte zu einer problematischen Beeinflussung der Gärkinetik.

Multi-omics Ansätze könnten in der Zukunft, angewandt werden um biologische Stoffwechselwege, die zu spezifischen Weinaromen führen, zu identifizieren, oder um den Einfluss spezifischer oenologischer Bedingungen zu untersuchen. Diese Ergebnisse sind nicht nur relevant für die Weinwirtschaft, sondern auch für andere Lebensmittelbranchen, in denen komplexe mikrobiologische Netzwerke relevant sind für die Herstellung fermentierter Lebensmittel, aber auch dem mikrobiellen Verderb von Roh- und Fertigware. Die in dieser Doktorarbeit präsentierten Ansätze könnten somit über den Wein hinaus in der Lebensmittelwirtschaft Anwendung finden.

Introduction

Returning to the cradle of wine fermentations

Microbial processes in food have accompanied human civilisations thousands of years. One such example is wine (Cavalieri et al., 2003; McGovern, 2013). Wine and other luxury beverages, such as coffee or whisky, are highly appreciated due to their complex and captivating aroma profiles (Bel-Rhlid et al., 2018). These properties are found to be a key dimension in wine quality (Charters and Pettigrew, 2007). Grape chemical composition and microbes, especially the yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) and bacteria *Oenococcus oeni* (*O. oeni*), play principal parts in generating the mixtures volatile compounds responsible for these aroma profiles.

Often, production processes need to be carefully controlled and managed if the aim is to fulfil certain quality requirements with a stable and routine production (Beuchat, 1993). With scientific researches, insight has been gained about the actual processes allowing these to have become more controllable. In wine, a prime example is development and introduction of yeast cultures. The origin of winemaking has been suggested to have involved native microbes carrying out alcoholic fermentations spontaneously (Chambers and Pretorius, 2010). However, following the discovery of yeast, Pasteur has been noted to study the “diseases of wine” (Barnett, 2000). Later spontaneous fermentations, which are here defined as fermentations in which no deliberate additions of microbes were done, was replaced by yeast cultures that were found to reduce many unwanted sensory characteristics and allow better fermentation reliability for practical purposes (Kraus et al., 1984; Rankine and Lloyd, 1963).

However, while the development of starter cultures as an alternative to spontaneous fermentation was triumphant, these cultures did not actually solve the problem of unreliable wine quality created during fermentation but rather circumvented it. Besides, the other main problem of spontaneous fermentations, the sluggish and stuck fermentations, remained a problem especially in white wines (Ribereau Gayon, 1999). The huge success of *S. cerevisiae* in research and coincident with the interest in commercial cultures, resulted in less emphasis placed on studies about non-

Saccharomyces yeasts (Kurtzman et al., 2011). Therefore the underlying potential of the nonconventional yeasts have remained unexploited but also the knowledge how wine was made thousands of year remains unexplored.

Recently, however, a multitude of incentives from various sources give renewed possibilities to return to these questions. With major experimental and analytical advancements, food research is becoming more capable in profiling and identifying the microbes and their functions in complex ecosystems, as well as how they are affected by different environmental conditions (Ferrocino and Cocolin, 2017). Thus, allowing to gather more information on interactions and other mechanisms such as in microbial diversity. Meanwhile the scientific evidence increasingly links health to the positive perception of food aroma compounds, because for example in tomatoes most relevant flavors derive from beneficial nutrients (Goff and Klee, 2006), consumers have started emphasizing food naturalness (Román et al., 2017), which is defined by no or low processing and supplementations. The food industry is tackling this with investigations emphasizing natural products and their processes (Bel-Rhlid et al., 2018). Therefore, advancing science and addressing the consumer calls for more naturalness are possible in unison.

One way to increase mutually beneficial new knowledge is returning the focus on basic but cornerstone questions that had been pushed aside previously. Many early wine research questions relate to spontaneous fermentation. Non-conventional yeasts and yeast interactions, in general, are suggested to increase wine aroma complexity thus positively influencing wine quality (Tempère et al., 2018). However, in both wine research and winemaking, a plethora of misconceptions exists, for example, about the origin and role of microbes, their functions and interactions. Some winemakers still see the practice of spontaneous fermentations as a way to achieve stylistic distinction (Pretorius, 2000), as well as exclusively allowing native yeasts can also be a requirement[1].

[1] 'No imported aromatic yeast or malo-lactic [sic] bacteria is permitted.' (Demeter Winemaking Standard for "Biodynamic Wine) https://www.demeter-usa.org/downloads/Demeter-Processing-Standards.pdf

Brief history of spontaneous and inoculated fermentations

While bacteria are important in some wine styles, yeasts play the major role for being solely responsible in transformation grape sugar to wine alcohol. Wine microbial research can be considered to begin when the main microbial species, *S. cerevisiae,* was isolated from vineyards (Pasteur, 1872). This led quickly to research about microbial cultures not only in wine and other fermented beverages (Chambers and Pretorius, 2010) but also allowed *S. cerevisiae* to become and remain the ideal experimental organism for modern biology (Botstein and Fink, 1988, 2011). Considering history of wine, the inoculation with microbial cultures has been relative quickly absorbed in modern winemaking.

The main problems of spontaneous fermentations were both reduction and oxidation of wine. As reviewed by Kraus et al. (1984), spontaneous fermentations were found to produce high amounts of SO_2 (70 mg/L) as well as rotten egg smelling hydrogen sulfide, but also white wine browning, acetaldehyde and ethyl acetate that associate with oxidation odours such as bruised apple and nail polish. Commercial cultures were introduced to winemaking 1965 and became gradually accepted in ensuring homogeneity of the final product (Pretorius, 2000; Reed and Nagodawithana, 1988). Inoculation of must with commercial culture cured many aroma defects, and lowered the incidence of sluggish fermentation, and thus became the standard in commercial winemaking.

However, arguments about style remained unresolved. While some winemakers resisted inoculations by claiming that distinctive wine styles were allowed by vineyard specific yeasts (Pretorius, 2000) whereas commercial yeasts were observed to mainly reduce volatile acidity with other effects on style remained more obscure and debated (Reed and Nagodawithana, 1988). Recently the use of commercial cultures has been suggested be some authors to have led to the sensorial standardisation of wine styles (Vigentini et al., 2016). Multiple studies have linked the roles of non-*Saccharomyces* yeasts in increasing sensory complexity (Reed and Nagodawithana, 1988; Tempère et al., 2018) and more specifically allowing the modulation of particular sensory characteristics to create unique wine style (Swiegers et al., 2005). Since the sequencing of *S. cerevisiae* genome (Goffeau et al., 1996), a general observation based on genetics and domestication is that most of industrial strains originated from only a few ancestral strains (Gallone et al., 2016), and

the genetic similarity of *S. cerevisiae* wine isolates echoes this observations (Borneman et al., 2016), thus supporting the sensory observations. However, while evidence is accumulating to microbial regionality (Bokulich et al., 2014; Knight et al., 2015), it has long been suggested that this microbial diversity might continue to play a role even after the inoculation with cultures (Reed and Nagodawithana, 1988).

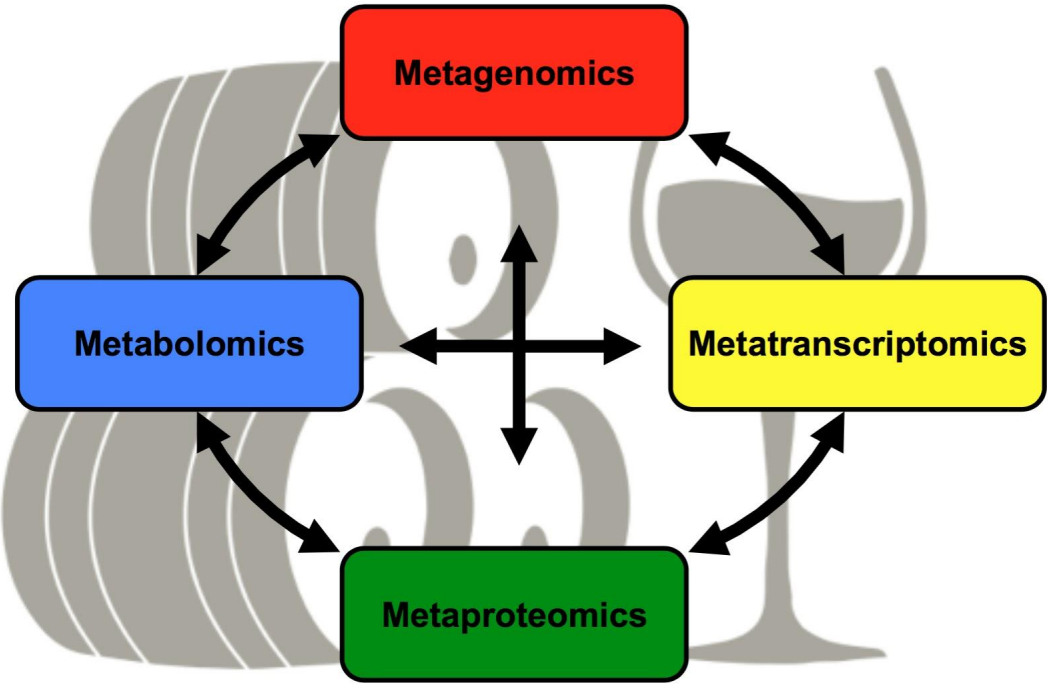
Technological and analytical advancements

Recent advances in high-throughput technologies have led to breakthrough approaches in which vast amount of data is generated, enabling a more in-depth understanding of the sample with a relative fast turnaround. These approaches are the so-called 'omics methods that all recover constituents directly from the sample, without traditional intermediary stages, such as isolation by culturing (Figure 1). While the original success of such approaches originates from marine environmental studies (Venter et al., 2004), they become a standard for many other microbiome international collaborations such as the human microbiome project[2] and the metagenomics and metadesign of the subways and urban biomes, MetaSUB[3].

For food-related studies, more recent examples of achievements of relevant 'omics approaches range from proteins (metaproteomics) (Hettich et al., 2013) and small molecules (metabolomics) (Bonner and Hopfgartner, 2018; Patti et al., 2012) to genetic data (De Filippis et al., 2018). Furthermore, the nucleic acid data can be RNA (metatranscriptomics) or DNA (metagenomics) and is often generated by using next-generation sequencing (NGS) that is associated with high-throughput sequencing (HTS) technologies. For DNA, where the non-targeted approach, shotgun sequencing, is considered true 'metagenomics' (De Filippis et al., 2018), research can also focus on a particular genetic regions of interest which is termed as 'metabarcoding' or amplicon-based sequencing.

2 https://hmpdacc.org/

3 http://metasub.org



.

Figure 1. An overview of linkages among multi-omics approaches in wine studies (Sirén et al., 2019b). Four 'omics approaches involve DNA (metagenomics), RNA (metatranscriptomics), proteins (metaproteomics) and metabolites (metabolomics). The application of more than one approaches is called multi-omics.

While traditional studies have focused on isolation by culturings, non-targeted studies, such as Denaturant Gradient Gel Electrophoresis (DGGE) and more recently NGS-based tools, allow the documentation of the microbial community composition. Non-targeted studies are needed as pure culture isolation and observation using artificial media is well-known to often yield incomplete understanding of a particular microbial flora (Henrici, 1933). In addition, many of subdominant microbes that may be viable may not be cultivated (Amann et al., 1995). In the development of increasingly economical NGS technologies, there is a growing interest in using these techniques in profiling microbial communities during wine fermentation and other food microbiology related studies. The studies applying metabarcoding to wine and food products are often focusing on amplifying bacterial 16S rRNA region and/or the ITS or D1/D2 26S regions for the fungi. Wine metabarcoding studies are relatively young field as reviewed by (Belda et al., 2017b; Stefanini and Cavalieri, 2018). First NGS-based wine study looked into the bacterial region (Bokulich et al., 2012); however, recent studies are often built on the fungal regions - given the importance of yeast in wine.

Currently, only two metagenomic wine studies have been published (Sternes et al., 2017;

Zepeda-Mendoza et al., 2018) as reviewed in Chapter 1 (Sirén et al., 2019b). One compares metagenomic and metabarcoding studies in wine fermentations (Sternes et al., 2017), while the other investigates interactions between lactic acid bacteria and the spoilage yeast *Brettanomyces bruxellensis* (Zepeda-Mendoza et al., 2018). Thus, there is considerable interest in profiling the whole genomic content using true metagenomic approaches, in the hope that they will also provide information about the functional pathways involved.

As the analytical technologies and methods are developing, the field is facing multiple challenges. One major challenge present in all 'omics approaches is the current poor state of annotation, whether identifying small molecules, genomes or their functions. Multiple solutions have been suggested for NGS, such as unsupervised database independent approaches, and utilising the sample size and other 'omic data in sequence assembly (Narayanasamy et al., 2016; Ye and Tang, 2016). For metabolomics, various annotation strategies for compound identification have been recently reviewed by two groups (Domingo-Almenara et al., 2018; Wolfender et al., 2018). In general, although the shortcomings of putative or partial identifications are acknowledged, they are essential and strategies to improve them continue to be developed. One such approach in metabolomics is relying on *in silico* fragmentation (Nguyen et al., 2018). Such approach and other fragmentation tools are actively developed to allow more metadata to be extracted from the individual feature (Blaženović et al., 2017). This also helps other approaches for exploring substructures untargeted (van der Hooft et al., 2016). Interestingly, also the database improvements are often composed from *in silico* fragmentation as well (Guijas et al., 2018). However, combining the two can easily been seen to potentially lead to fallacious interpretations, needing proper tandem mass spectra data from one or the other. Yet, integrating novel data analysis with improvements in discovery based approaches such as data-independent acquisition (DIA) analysis (Doerr, 2014) are generating new possibilities while delaying the need for precise compound annotation.

Nevertheless, relatively little has changed for the core idea of the metabolomic approach (Viant et al., 2017). In general, the highest level of confidence of identifying metabolites needs chemical reference standards and orthogonal analytical techniques are often impractical, thus leading to investigators to often focus on only database identified peaks

(Viant et al., 2017). Thus metabolomics is generally concluded to be surrounded by dark matter (Peisl et al., 2018; Schymanski et al., 2017). In general, the various approaches to improve the quality of ‘omic databases are needed, as also NGS databases are often consisting of poorly annotated genes and genomes (Belda et al., 2017a; Sczyrba et al., 2017). Novel approaches, such as open-source competitions where data tool limits are realised, can impact by changing the field rapidly (Schymanski et al., 2017; Sczyrba et al., 2017). Additionally, further information can be established from other studies by using ontologies (Gruber, 1995), such as chemical “taxonomies” (Djoumbou Feunang et al., 2016) or orthologous groups of proteins (Huerta-Cepas et al., 2016).

By using modern bioinformatic approaches, such as network theory and statistical learning, it is possible to investigate both biological components and to determine their interactions and functions (Barabási and Oltvai, 2004). This can be done by combining different ‘omics methods, coined as multi-omics. The top-down network view is known to be useful when investigating biological systems (Bray, 2003). One promising method for combining multi-omic data was recently developed allowing to determine topology of relationships between different multiple datasets (Aben et al., 2018).

Another crucial challenge of ‘omics approaches is the lack of standardisation to allow research reproducibility and inter-study comparisons (Knight et al., 2018a). The standardisations not only rely on critical experimental method comparisons such as DNA extraction methods, but also generate more transparency in data processing. This transparency can be further increased by allowing reproducibility possibilities through open-access and automatised computational workflows (Kang et al., 2015; Sirén et al., 2019a) which themselves can be integrated into bigger architectures (Murat Eren et al., 2015; Peters et al., 2018).

Albeit the multiple challenges, multi-omic data integration approaches have been showing promise in systems biology studies such as ecology studies (Friedman and Gore, 2017) and protein interactions (Aloy and Russell, 2006). Therefore these approaches can be further applied to wine and fermentation process which are also inherently biological systems. While still limited applications exist in wine research, applying these approaches offers new venues to advance our understanding of wine microbial communities and their potential functions and the sensory-chemistry interface. These multi-omic approaches can

be also of great interest to other food industries where complex microbial networks have an essential impact.

Riesling characteristics

Vitis vinifera cultivars (cv.) have long been appreciated having specific sensory properties across cultivars and regions. These specificities were already noted in AD 77 (Robinson and Harding, 2015). Although in this sense, Riesling (*Vitis vinifera* cv.) is relatively young cultivar with the first recorded whole vineyard planted in 1720[4]. Riesling is considered to be one of the noble cultivar due to its' capability of producing geographically distinct (Bauer et al., 2011; Fischer et al., 2016) and age-able wines (Simpson and Miller, 1983). This nobleness continues to be confirmed by consumers, as 10 out of 50 the world's most expensive wines are Riesling[5]. Alas, more marketing could be done as no Riesling wines are currently on the top 100 of most searched wines[6].

Wines made from Riesling have long been known to deviate recognisably from other wines based on their sensory properties (Winton et al., 1975). These can range from fruity odour with citrus (lemon/grapefruit), peach and exotic fruits (mango and passion fruit), all the way to honey, floral, smoky, green and mineral odours (Bauer et al., 2011; Schüttler et al., 2015). Additionally, its characteristic varietal odour, kerosene, can sometimes be too high, therefore considered actually impairing the perceived regional typicality (Schüttler et al., 2015).

Geological and geographical factors have been shown to affect sensory properties of Riesling, for example associating particular aroma notes for bedrock types (Bauer et al., 2011; Fischer et al., 2016). On the other hand, observations have been made on the role of yeasts in producing different aroma profiles (Benito et al., 2015; Egli et al., 1998; Fischer, 2007). It has been concluded that different inoculated yeasts produced different aroma and in general increased fruit profile (Benito et al., 2015; Fischer, 2007). However, the mechanisms and the relation to spontaneous fermentations are unclear. Also, the use of

4 https://www.schloss-johannisberg.de/en/history.htm

5 https://www.wine-searcher.com/most-expensive-wines (List last updated: 2nd January 2019)

6 https://www.wine-searcher.com/find/- (List last updated: 2nd January 2019)

standard winemaking tools such as SO_2 or nutritional supplements often further complicates the matter (Egli et al., 1998).

Therefore, the general sensorial outcome of the microbes and especially their interactions remains inconclusive with both positive and negative effects possible (Masneuf-Pomarede et al., 2015), as inoculation based studies will not infer a conclusive outcome. Microbial interactions are vital to the winemaking process with numerous different microbes known to be involved in the formation of wine flavour and aroma. Recent studies have shed light upon the roles of microbial diversity in vineyards (Bokulich et al., 2014, 2016; Morrison-Whittle and Goddard, 2018) and wineries (Bokulich et al., 2013; Ganucci et al., 2018; Stefanini et al., 2016). Understanding the communities and their interaction throughout the process stands to enhance our knowledge of winemaking and wine complexity (Tempère et al., 2018).

Wine aroma and sensory challenges

The chemical composition of grape must plays a significant role in formation of the sensory perception of wine aroma. Wine aroma is composed of sensory relevant volatile compounds or odourants. The compositions of aroma, odourants and grape must constituents are all known to vary along the climate, viticulture management, cultivar and vintage. As noted above, microbes, especially yeasts, impact various ways on wine aroma formation. Below, this impact is briefly examined with focus placed on Riesling varietal aroma compounds that can be influenced by yeasts, the most important microbes during alcoholic fermentation. For more detailed reviews, see recent extensive reviews on yeast aroma formation (Dzialo et al., 2017) and overall microbial contribution to wine aroma (Belda et al., 2017a).

Hundreds of volatile aroma compounds have been identified in wine, and these compounds are thought to derive from multiple sources: grapes and processing, fermentation and ageing (Fischer, 2007). Table 1 lists several important aroma compounds, which have been found in Riesling wines, but are also commonly found in beer. These listed aromas constitute 10% of the main odourants in any food item (Dunkel et al., 2014). This group consists of fermentation related ethyl esters and higher alcohol acetates, but also varietal aroma compounds such as monoterpene alcohols, C_{13}-

norisoprenoids, polyfunctional thiols and volatile phenols (original compound identifications in Riesling refer to synopsis from Schüttler (Schüttler, 2012; Schüttler et al., 2015).

While *S. cerevisiae* is the main producer of volatile compounds in all fermented beverages, other yeast species also have an influence. Wine studies about the influence of non-*Saccharomyces* inoculations have been recently reviewed (Varela, 2016). This influence can be unpredictable as the results often depend on the grape must composition (Holt et al., 2018). There is also an indirect influence on volatile compounds originating from microbial regionality. Although further research is clearly needed, initial hypotheses have been made linking the microbial and volatile compositions. The genetic variance of *S. cerevisiae* in microbial populations was suggested to be altering wine phenotypes as observed with *S. cerevisiae* (Knight et al., 2015) and, in general, bacteria and fungi communities were found to correlate with regional metabolite profiles (Bokulich et al., 2016).

Fermentation aromas

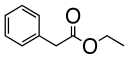
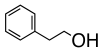
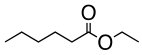
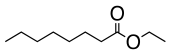
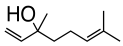
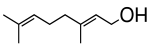
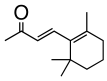
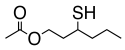
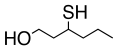
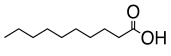
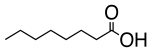
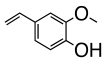
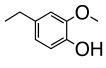
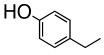
Saccharomyces sp. has often been used as model species for studying formation of fermentation derived volatile compounds, higher alcohols and esters, the main compounds found in in wine fermentations. Examples of such studies are a study of formation of higher alcohols via Ehrlich pathway, or the discovery of enzymes responsible for the formation of esters (Gamero et al., 2016). However, these observations are not always applicable to non-*Saccharomyces* yeasts because of metabolic differences for example in the regulation of Ehrlich pathway (Gamero et al., 2016). Differences regarding fermentation volatile compound formation can be either strain-related or species-related, although the exact mechanisms are not understood yet (Gamero et al., 2016).

First main group of fermentation origin, the higher alcohols such as 2-phenylethanol or isobutanol, result from amino acid metabolism by yeasts as reviewed by Michel et al. (2016). They are in general considered as negative but research remains non-unanimous with complicated contributors to wine aroma from ranging suppressive (methionol), green (isoamyl alcohol (Aznar et al., 2003; Sáenz-Navajas et al., 2018)) to rose (2-phenylethanol (Katarína et al., 2014)). Furthermore, isobutanol-isoamyl alcohol pair has

been suggested to play a sensorial impact in a red wine study (de-la-Fuente-Blanco et al., 2016); however, others emphasise positive aromas while acknowledging that higher doses can lead to a negative influence (Capozzi et al., 2015) and have been suggested to contribute to the wine atypical ageing odour (Fischer, 2007).

The second main fermentation group, esters, have been suggested to have the major role in wine aroma as they are the most abundant volatiles in the matrix (Nykänen, 1986). Despite being abundant in wine and other fermented beverages, these fermentation aroma compounds might be considered affecting varietal aroma indirectly with microbial regionality, although compounds were not specified (Bokulich et al., 2016). Moreover, some esters such as ethyl cinnamate have been suggested to play a direct role in varietal aroma (Loscos et al., 2007; Moio and Etievant, 1995). While amino acid composition clearly affecting the volatile compound composition in wine, further insights could be gained by improving analytical and statistical methods (Hernández-Orte et al., 2005).

Table 1. Common aroma compounds and families found in food. These compounds are found in Riesling (Fischer, 2007; Schüttler et al., 2015), beer (Michel et al., 2016; Zheng et al., 2017), and food items in general (Dunkel et al., 2014). A % = relative odor activity value based abundance in food products (Dunkel et al., 2014).

Compound family	Name	A %	Structure
Acetate esters	isoamyl acetate	13,7	
	ethyl acetate	1,3	
	ethyl 2-phenylacetate	0,9	
Carbonyl compounds	diacetyl	41,9	
	acetaldehyde	28,6	
Higher alcohols	2−phenylethanol	22,9	
	isoamyl alcohol	13,7	
	isobutanol	2,6	
Medium-chain fatty acid ethyl esters	ethyl hexanoate	16,3	
	ethyl octanoate	9,3	
Monoterpene alcohols	linalool	24,7	
	geraniol	4,8	
Norisoprenoids	β-damascenone	24,2	
	β-ionone	6,2	
Thiols	3-sulfanylhexyl acetate	2,2	
	3-sulfanylhexan-1-ol	2,2	
Volatile organic acids	acetic acid	29,1	
	decanoic acid	4,4	
	octanoic acid	3,5	
Volatile phenols	4−vinylguaiacol	13,7	
	4−ethylguaiacol	11	
	4−ethylphenol	3,1	

Other fermentation related volatile compounds are sulphuric compounds, volatile organic acids, Strecker aldehydes and other carbonyl compounds. In synthetic media and beer, the amount of organic acids has been found to increase in relation to fermentation time while the compound composition was found varying (Coote and Kirsop, 1974). Yeast central metabolism plays a key role in acid levels since most acids originated as byproducts of glycolysis, citric acid cycle, amino acids and fatty acid metabolism (Michel et al., 2016). The chemical and sensory significances of aldehydes have previously been studied in beer (Baert et al., 2012) for their negative sensory property, (staleness), and recently also in wine, such as methional, acetaldehyde and 2-phenylacetaldehyde (Bueno et al., 2018). In beer, yeasts have been found to be able to excrete Strecker aldehydes during fermentation albeit with limited sensory influence (Baert et al., 2012).

Varietal aroma

The main varietal aroma in Riesling originates from a specific combination of volatile compounds (Fischer, 2007; Rapp and Mandery, 1986). The volatile compounds, such as monoterpenes, C13-norisoprenoids and polyfunctional thiols, are often present in grape must as non-volatile precursors, for example as glycosides or S-conjugates, which can be subsequently released during fermentation or ageing by different microbial activities. Additionally the liberation can happen via acid hydrolysis, but also transformation of both the glycosidic and the free form.

Some of these varietal volatile compounds, such as the monoterpenes linalool and geraniol, can be found both free and as glycosides in grapes (Williams et al., 1982). Besides short-term goals of maximising aroma release by winemaking, glycosidic precursors might play a role of sensory important as well during wine aging as reviewed by Marais (1983). Additionally, once liberated, the ratio of these compounds can differ as they can be transformed by acid catalysed isomerisation or enzymatic reactions (King and Dickinson, 2000). It has been observed that cultivar influences this composition (Skinkis et al., 2008) and yeasts play a major role in building it (Delfini et al., 2001; Loscos et al., 2007), although the transformation is directed mainly to the formation of linalool and α−terpineol from their glucosides (King and Dickinson, 2000). Furthermore some microbes and S. cerevisiae are also able to produce terpenoids de novo (Carrau et al., 2005; King and Dickinson, 2000), which further complicates understanding of aroma origins.

Besides monoterpenes, other glycosylated precursors are C_{13}-norisoprenoids, linear alcohols and benzene derivatives such as phenols (Sarry and Günata, 2004). Non-*Saccharomyces* yeasts have been found to be able to release glycosides more effectively than *S. cerevisiae* (Zoecklein et al., 1997). In additional to glucose and fructose, non-*Saccharomyces* yeasts can utilise various kinds of sugars, such as lactose, galactose, raffinose, sucrose, trehalose and maltose (Michel et al., 2016). This highlights possibilities for discovering novel metabolic pathways as well as new possible enzymatic activities in releasing the aglycone moieties from their glycosidic precursors, which might be important also for wine aroma profile

C_{13}-Norisoprenoids are a diverse group of compounds that are likely originating from carotenoid degradation, but often can result from more complex chemical rearrangements (Winterhalter et al., 1990). Most important norisoprenoids in Riesling are 1,1,6-Trimethyl-1,2-dihydronaphthalene (TDN) and ß-damascenone (Winterhalter and Gök, 2013). The contribution of β-damascenone to wine aroma has been commented numerous times but with its important impact remaining vague, with potential impact as an enhancer of fruity-type of aroma intensities (Sefton et al., 2011). However, the microbial influence on these compounds is less well established (Swiegers et al., 2005; Winterhalter and Gök, 2013).

In beer and wine, volatile phenols have been traditionally regarded as an issue originating from *Brettanomyces* spp., but currently more emphasis is spent on them due to the increase in oenological problems of smoke-tainted wines (Kennison et al., 2007). Volatile phenols are also found in their glycosylated forms (Hayasaka et al., 2010; Noestheden et al., 2018). While yeast-produced phenols can be referred as off-flavour in most beer styles (Michel et al., 2016), some cultivars such as Riesling, Gewürztraminer and especially Koshu (*V. vinifera* x *V. davidii*) (Goto-Yamamoto et al., 2015) have high amounts of hydroxycinnamic acids (Nagel et al., 1979; Okamura and Watanabe, 1981), therefore with the current emphasis to volatile phenols, the hydroxycinnamic acid contributions to varietal aroma or flavor might be reinvestigated (Moio and Etievant, 1995; Vèrette et al., 1988).

Polyfunctional thiols are the most positively influencing sulphuric compounds on wine aroma (2011; Swiegers et al., 2005). Yeast activity is required to release polyfunctional

thiols from their cysteinylated and glutathionylated precursor forms (Dubourdieu et al., 2006; Tominaga et al., 2000) although grape processing can play a huge role for their liberation but also formation (Allen et al., 2011; Capone et al., 2012; Capone and Jeffery, 2011; Maggu et al., 2007). Yeast activity and fungal diversity in general have been found to significantly modulate the formation of thiols although more research is still needed to understand how they actually relate to them (Anfang et al., 2009; Deroite et al., 2018; Dubourdieu et al., 2006; Knight et al., 2018b; Ruiz et al., 2018).

Sensory challenges

Two main challenges exist in linking wine aroma compounds to actual effect on sensory properties. First challenge underscores the analytical and experimental design. While new possibilities with the advancements in technology are being showcased by identifying new glycosylated volatile phenols and monoterpenes in wine (Hjelmeland et al., 2015; Noestheden et al., 2018), further advances in mass spectrometry could shed more light upon the inherent structural diversity of plant metabolites in general (Hofmann et al., 2015). Non-targeted approaches can be extremely powerful for screening as demonstrated for wine (Hranilovic et al., 2018). Open science based solutions such as open databases and network tools in Global Natural Products Social Molecular Networking (GNPS) (Wang et al., 2016), are being actively developed, as database issues are increasingly becoming a bottleneck for such non-targeted studies. Semi-targeted approaches are however still needed due to instrumental sensitivity and selectivity issues, as relevant compounds found in wine can differ more than 10^9-fold. Additionally, in echoing Tukey (1980), it needs to be stressed that targeted studies are not enough to find novel patterns, although non-targeted studies need to be complimented with targeted confirmatory studies, such as confident compound identifications.

Secondly, issues establishing links between chemical composition of wine and sensory properties need to be addressed. Similar to wine, in other fermented beverages such as cachaça, yeasts may influence the chemical composition of the final product (Portugal et al., 2017) and differences correlate with geographical origin (Serafim et al., 2016). It seems that a potential link might exist between chemicals and sensory properties (Serafim et al., 2013). However, these relationships remain vague at most. While wine research is increasingly using multivariate tools, it still continues to use methods relying on

(multi)linear relationships to solve the link between sensory and chemistry such as partial least-squares (PLS) regression (Fischer et al., 1996; Lee and Noble, 2003, 2006; Noble and Ebeler, 2002; Sáenz-Navajas et al., 2018; Schüttler et al., 2015; Siebert et al., 2018). Nevertheless, these methods have allowed new observations about the multivariate nature of sensory perceptions, such as apricot odour (Siebert et al., 2018) or a green characteristic in red wine (Sáenz-Navajas et al., 2018) that is not methoxypyrazine-related (Allen et al., 1991).

However there is a strong need to apply more potent new statistical tools as the relationship between sensory perception and the chemical concentration might not be linear even in multivariate space. In order to make it linear in univariate space, the chemical concentrations need to be transformed into corresponding odour activity values (OAV) which are defined from odour thresholds. These OAV further need then to be standardised for linearity estimations of actual perceived odour intensity (Wu et al., 2016). While OAV can be seen giving more information about the actual sensorial effects than chemical concentrations, these values should be treated with care (Wu et al., 2016). The determination of sensory thresholds is known to be heavily influenced by the actual wine matrix (Fischer, 2007), but also to be a significant effect since it is inherently an estimation, a confidence has to be set (Lawless and Heymann, 2010). This is crucial since the values are based on a model that acts in logarithmic space, meaning that the errors are also inverse to that space and become exponential. Although OAV are still considered important, the underlying issue has to be solved to be more applicable for comparable and reproducible studies.

Interestingly, it was recently suggested that only 230 food odourants were found important for food in general (Dunkel et al., 2014) while at the same time human is suggested to have the capability to discriminate more than one trillion of smells (Bushdid et al., 2014) which are mainly originating from mixtures of a subset of 128 of these odourants. Therefore, new statistical approaches are also required not only for studying the interaction of these odourants, but also to reveal the actual sensory relevant correlations. One approach for replacing OAV could be relating the chemical importance to concentration–response relations of olfactory receptors (Geithe et al., 2017). However, another approach could be investigating OAV as a proportional value of the sample aroma, instead of an absolute intensity score. This would allow usage of modern

ecological statistical tools developed for similar problems. Nevertheless, biological responses also play a role (Poivet et al. 2018).

Structure, aims and contributions

In order to make high quality wine, one prevailing approach is spontaneous fermentation. However, many of winemaking decisions are based on tradition instead of scientific evidence. Therefore, investigating spontaneous fermentations, native yeasts and their interactions lays the ground for this thesis. The aim of this thesis is to study the role of the microbial communities from vineyard during spontaneous fermentation by using distinct modern approaches and techniques. Throughout my PhD, I have gained new knowledge about how to examine the impact originating from microbial diversity of spontaneous wine fermentations by applying targeted and non-targeted approaches for chemical and biological factors. Firstly, the role of multiple 'omics approaches is evaluated with emphasis on the future potential of such approaches for the wine community. Following, the feasibility and benefit of new state-of-the art techniques for mining non-targeted GC-MS data are illustrated by using a novel and fully automated machine learning approach. Further focus is placed on reproducibility by standardisation 'omics, namely microbial analyses using both metabarcoding and metagenomics approaches as well as targeted GC-MS and sensory analysis. Riesling was used as the basis, for three out of five chapters. In order to achieve proper comparisons between controlled experimental winemaking and participating wine estates' customary vinifications, I refined pilot-scale and microvinification methods to recover the microbial diversity present during spontaneous fermentation. Besides understanding and improving high-wine quality, the main motivation of this PhD was to contribute to reproducibility and open science.

The thesis is organised in the following chapters: Chapter 1 sets the stage by addressing recent developments and potential in 'omics approaches for the wine community. Chapters 2 and 3 are method development sections utilising multiple 'omics datasets. In Chapter 4 the workflow is applied by using exploratory analysis and it acts as a proof-of-concept study for Chapter 5. Chapter 5 investigates the effects on chemical composition and sensory properties originating from spontaneous fermentations with interacting conventional way between microbial communities in both vineyard and cellar, and only looking at the effect rising from vineyard microbial communities using novel statistical approaches.

Chapter 1

Chapter 1 is entitled “Multi-omics and potential applications in wine production”. This review manuscript was published online in Current Opinion in Biotechnology on December 2018 as part of a special issue on Food Biotechnology to be published in April 2019. It was co-authored with Sarah Siu Tze Mak, Ulrich Fischer, Lars Hestbjerg Hansen and Tom Gilbert. This review describes multiple ‘omics approaches in wine giving an emphasis on potential applications and challenges to wine community, with also adding future perspective of combining multi-omics approach that were further developed in chapter 5 and utilised in ongoing wine-related projects (not included in the thesis). My contribution included the conception and design of the review, literature reviews and co-writing the manuscript.

Chapter 2

Chapter 2 is entitled “Automated supervised learning pipeline for non-targeted GC-MS data analysis”. This full length article manuscript was accepted and published in Analytica Chimica Acta: X on January 2019 in the first volume of the online journal. It was co-authored with Ulrich Fischer and Jochen Vestner. This manuscript reports the development of a novel, fully automated analysis of non-targeted GC-MS data. We show feasibility and benefits of tensor decomposition of segmented GC-MS chromatograms prior to supervised classification of samples using a machine learning pipeline. The new approach was applied to three different datasets from recently published studies and the performance of our approach was compared with the results of the original publications. I conceived the study and contributed to experimental design, programming, acquisition, analysis and interpretation of the data and writing the manuscript.

Chapter 3

Chapters 3 is entitled “Comparison of DNA extraction methods for use in fungal diversity analyses in Riesling during alcoholic fermentation”. This manuscript was submitted to International Journal of Food Microbiology for review in January 2019. The study is co-authored by Sarah Siu Tze Mak, Christian Carøe, Shyam Gopalakrishnan, Martin René Ellegaard, Franziska Klincke, Anders J. Hansen, Ulrich Fischer and Tom Gilbert. This

manuscript shows that DNA extraction method has to be standardized for reliable HTS-studies in wine. As the amount of wine studies using HTS-based methods continue to increase, studies from other fields point out that available DNA extraction methods vary. In order to increase data reproducibility in prospective wine HTS-based studies, three different DNA extraction methods were compared using native Riesling fermentations from different vineyards. Significant differences were observed with extracted DNA amount, and number of gene copies but also with ease of use and handling time. By using metabarcoding it was observed that all three methods could distinguish the effect of the vineyard on fungal diversity but with quantitative differences. This highlights the importance of standardizing DNA extraction methods for wine HTS-based studies. The method that was observed the fastest and gave the highest fungal diversity, was applied for DNA extraction in Chapter 4 (and other ongoing projects not included in this thesis). My contributions included experimental design, sample processing, data analysis, data interpretation and co-writing the manuscript.

Chapter 4

Chapter 4 is entitled: "Taxonomic and functional characterisation of the microbial community during spontaneous in vitro fermentation of Riesling must". This is co-authored by Sarah Siu Tze Mak, Chrats Melkonian, Christian Carøe, Jan Hendrik Swiegers, Douwe Molenaar, Ulrich Fischer and Tom Gilbert. This manuscript was submitted to Frontiers in Microbiology in January 2019. This manuscript shows that metagenomic methods allow both taxonomic and functional conclusions to be derived from microbial wine studies revealing new patterns. Since confusion exists about the ability of vineyard microbes to carry out wine fermentations, a Riesling in vitro spontaneous fermentation study, including only vineyard microbes, was carried out under sterile conditions. Results reveal that sufficient amount of Saccharomyces is present in the vineyard as the wines fermented to dryness, however differing in required time. Both metabarcoding and metagenomic approaches were used for investigating the microbial community composition. Additionally, metagenomic approaches together with functional analysis in wine studies were applied to investigate functional potentials of vineyard microbes. We observed novel bacterial growth during alcohol fermentation, and could hypothesize a relevant pathway of how non-Saccharomyces yeast interacts with bacteria. My contributions included experimental design, data generation, data analyses and

interpretations, curating the wine metagenomic database, and co-writing the manuscript.

Chapter 5

Chapters 5 is entitled: “Sensory relevant chemical interactions revealed with Riesling wines fermented with different microbial communities”. This is co-authored by Inês Oliveira, Vicente Ferreira and Ulrich Fischer. The manuscript is a in preparation draft to be submitted to Food Research International. This original research applies the exploratory data analysis and Bayesian statistical approaches to investigate the effects of population interactions in wine fermentations. It is shown that grossing chemical interaction studies can generate also sensory relevant information. Wine chemical and sensorial variability was shown to be increased with vineyard-winery interactions, thus showing that such studies also benefit for integrating different ‘omics approaches. My contributions included experimental design, experimental winemaking, data generation, data analyses and interpretations, writing the manuscript draft.

References

Aben, N., Westerhuis, J. A., Song, Y., Kiers, H. A. L., Michaut, M., Smilde, A. K., et al. (2018). iTOP: inferring the topology of omics data. *Bioinformatics* 34, i988–i996.

Allen, M. S., Lacey, M. J., Harris, R. L. N., and Vance Brown, W. (1991). Contribution of Methoxypyrazines to Sauvignon blanc Wine Aroma. *Am. J. Enol. Vitic.* 42, 109–112.

Allen, T., Herbst-Johnstone, M., Girault, M., Butler, P., Logan, G., Jouanneau, S., et al. (2011). Influence of grape-harvesting steps on varietal thiol aromas in Sauvignon blanc wines. *J. Agric. Food Chem.* 59, 10641–10650.

Aloy, P., and Russell, R. B. (2006). Structural systems biology: modelling protein interactions. *Nat. Rev. Mol. Cell Biol.* 7, 188–197.

Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.

Anfang, N., Brajkovich, M., and Goddard, M. R. (2009). Co-fermentation with Pichia kluyveri increases varietal thiol concentrations in Sauvignon Blanc. *Aust. J. Grape Wine Res.* 15, 1–8.

Aznar, M., López, R., Cacho, J., and Ferreira, V. (2003). Prediction of aged red wine aroma properties from aroma chemical composition. Partial least squares regression models. *J. Agric. Food Chem.* 51, 2700–2707.

Baert, J. J., De Clippeleer, J., Hughes, P. S., De Cooman, L., and Aerts, G. (2012). On the origin of free and bound staling aldehydes in beer. *J. Agric. Food Chem.* 60, 11449–11472.

Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113.

Barnett, J. A. (2000). A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850-1880. *Yeast* 16, 755–771.

Bauer, A., Wolz, S., Schormann, A., and Fischer, U. (2011). 'Authentication of Different Terroirs of German Riesling Applying Sensory and Flavor Analysis', in *Progress in Authentication of Food and Wine* ACS Symposium Series. (American Chemical Society), 131–149.

Belda, I., Ruiz, J., Esteban-Fernández, A., Navascués, E., Marquina, D., Santos, A., et al. (2017a). Microbial Contribution to Wine Aroma and Its Intended Use for Wine Quality Improvement. *Molecules* 22, 189.

Belda, I., Zarraonaindia, I., Perisin, M., Palacios, A., and Acedo, A. (2017b). From Vineyard Soil to Wine Fermentation: Microbiome Approximations to Explain the 'terroir' Concept. *Front. Microbiol.* 8, 8966.

Bel-Rhlid, R., Berger, R. G., and Blank, I. (2018). Bio-mediated generation of food flavors – Towards sustainable flavor production inspired by nature. *Trends Food Sci. Technol.* 78, 134–143.

Benito, S., Hofmann, T., Laier, M., Lochbühler, B., Schüttler, A., Ebert, K., et al. (2015). Effect on quality and composition of Riesling wines fermented by sequential inoculation with non-Saccharomyces and Saccharomyces cerevisiae. *Eur. Food Res. Technol.* 241, 707–717.

Beuchat, L. R. (1993). Selective media for detecting and enumerating foodborne yeasts. *Int. J. Food Microbiol.* 19, 1–14.

Blaženović, I., Kind, T., Torbašinović, H., Obrenović, S., Mehta, S. S., Tsugawa, H., et al. (2017). Comprehensive comparison of in silico MS/MS fragmentation tools of the CASMI contest: database boosting is needed to achieve 93% accuracy. *J. Cheminform.* 9, 32.

Bokulich, N. A., Collins, T. S., Masarweh, C., Allen, G., Heymann, H., Ebeler, S. E., et al. (2016). Associations among Wine Grape Microbiome, Metabolome, and Fermentation Behavior Suggest Microbial Contribution to Regional Wine Characteristics. *MBio* 7.

Bokulich, N. A., Joseph, C. M. L., Allen, G., Benson, A. K., and Mills, D. A. (2012). Next-generation sequencing reveals significant bacterial diversity of botrytized wine. *PLoS One* 7, e36357.

Bokulich, N. A., Ohta, M., Richardson, P. M., and Mills, D. A. (2013). Monitoring Seasonal Changes in Winery-Resident Microbiota. *PLoS One* 8, e66437.

Bokulich, N. A., Thorngate, J. H., Richardson, P. M., and Mills, D. A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc. Natl. Acad. Sci. U. S. A.* 111, E139–48.

Bonner, R., and Hopfgartner, G. (2018). SWATH data independent acquisition mass spectrometry for metabolomics. *Trends Analyt. Chem.* doi:10.1016/j.trac.2018.10.014.

Borneman, A. R., Forgan, A. H., Kolouchova, R., Fraser, J. A., and Schmidt, S. A. (2016). Whole Genome Comparison Reveals High Levels of Inbreeding and Strain Redundancy Across the Spectrum of Commercial Wine Strains of Saccharomyces cerevisiae. *G3* 6, 957–971.

Botstein, D., and Fink, G. R. (1988). Yeast: an experimental organism for modern biology. *Science* 240, 1439–1443.

Botstein, D., and Fink, G. R. (2011). Yeast: an experimental organism for 21st Century biology. *Genetics* 189, 695–704.

Bray, D. (2003). Molecular networks: the top-down view. *Science* 301, 1864–1865.

Bueno, M., Marrufo-Curtido, A., Carrascón, V., Fernández-Zurbano, P., Escudero, A., and Ferreira, V. (2018). Formation and Accumulation of Acetaldehyde and Strecker Aldehydes during Red Wine Oxidation. *Front Chem* 6, 20.

Bushdid, C., Magnasco, M. O., Vosshall, L. B., and Keller, A. (2014). Humans can discriminate more than 1 trillion olfactory stimuli. *Science* 343, 1370–1372.

Capone, D. L., Black, C. A., and Jeffery, D. W. (2012). Effects on 3-mercaptohexan-1-ol precursor concentrations from prolonged storage of Sauvignon blanc grapes prior to crushing and pressing. *J. Agric. Food Chem.* 60, 3515–3523.

Capone, D. L., and Jeffery, D. W. (2011). Effects of transporting and processing Sauvignon blanc grapes on 3-mercaptohexan-1-ol precursor concentrations. *J. Agric. Food Chem.* 59, 4659–4667.

Capozzi, V., Garofalo, C., Chiriatti, M. A., Grieco, F., and Spano, G. (2015). Microbial terroir and food innovation: The case of yeast biodiversity in wine. *Microbiol. Res.* 181, 75–83.

Carrau, F. M., Medina, K., Boido, E., Farina, L., Gaggero, C., Dellacassa, E., et al. (2005). De novo synthesis of monoterpenes by Saccharomyces cerevisiae wine yeasts. *FEMS Microbiol. Lett.* 243, 107–115.

Cavalieri, D., McGovern, P. E., Hartl, D. L., Mortimer, R., and Polsinelli, M. (2003). Evidence for S. cerevisiae fermentation in ancient wine. *J. Mol. Evol.* 57 Suppl 1, S226–32.

Chambers, P. J., and Pretorius, I. S. (2010). Fermenting knowledge: the history of winemaking, science and yeast research. *EMBO Rep.* 11, 914–920.

Charters, S., and Pettigrew, S. (2007). The dimensions of wine quality. *Food Qual. Prefer.* 18, 997–1007.

Coote, N., and Kirsop, B. H. (1974). The Content Of Some Organic Acids In Beer And Other Fermented Media. *J. Inst. Brew.* 80, 474–483.

De Filippis, F., Parente, E., and Ercolini, D. (2018). Recent Past, Present, and Future of the Food Microbiome. *Annu. Rev. Food Sci. Technol.* 9, 589–608.

de-la-Fuente-Blanco, A., Sáenz-Navajas, M.-P., and Ferreira, V. (2016). On the effects of higher alcohols on red wine aroma. *Food Chem.* 210, 107–114.

Delfini, C., Cocito, C., Bonino, M., Schellino, R., Gaia, P., and Baiocchi, C. (2001). Definitive Evidence for the Actual Contribution of Yeast in the Transformation of Neutral Precursors of Grape Aromas. *J. Agric. Food Chem.* 49, 5397–5408.

Deroite, A., Legras, J.-L., Rigou, P., Ortiz-Julien, A., and Dequin, S. (2018). Lipids modulate acetic acid and thiol final concentrations in wine during fermentation by Saccharomyces cerevisiae × Saccharomyces kudriavzevii hybrids. *AMB Express* 8, 130.

Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., et al. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminform.* 8, 61.

Doerr, A. (2014). DIA mass spectrometry. *Nat. Methods* 12, 35.

Domingo-Almenara, X., Montenegro-Burke, J. R., Benton, H. P., and Siuzdak, G. (2018). Annotation: A Computational Solution for Streamlining Metabolomics Analysis. *Anal. Chem.* 90, 480–489.

Dubourdieu, D., Tominaga, T., Masneuf, I., des Gachons, C. P., and Murat, M. L. (2006). The Role of Yeasts in Grape Flavor Development during Fermentation: The Example of Sauvignon blanc. *Am. J. Enol. Vitic.* 57, 81–88.

Dunkel, A., Steinhaus, M., Kotthoff, M., Nowak, B., Krautwurst, D., Schieberle, P., et al. (2014). Nature's Chemical Signatures in Human Olfaction: A Foodborne Perspective for Future Biotechnology. *Angew. Chem. Int. Ed.* 53, 7124–7143.

Dzialo, M. C., Park, R., Steensels, J., Lievens, B., and Verstrepen, K. J. (2017). Physiology, ecology and industrial applications of aroma formation in yeast. *FEMS Microbiol. Rev.* 41, S95–S128.

Egli, C. M., Edinger, W. D., Mitrakul, C. M., and Henick-Kling, T. (1998). Dynamics of indigenous and inoculated yeast populations and their effect on the sensory character of Riesling and Chardonnay wines. *J. Appl. Microbiol.* 85, 779–789.

Ferrocino, I., and Cocolin, L. (2017). Current perspectives in food-based studies exploiting multi-omics approaches. *Current Opinion in Food Science* 13, 10–15.

Fischer, U. (2007). 'Wine Aroma', in *Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability*, ed. R. G. Berger (Springer Berlin Heidelberg), 241–267.

Fischer U., Bauer, A., Koschinski S., and Schmarr, H.-G. (2016). Terroir in the Old and New World – what sensory is telling us. in *Proc. 16th Austr. Wine Ind. Tech. Conf.*, Adelaide, 24-28th July, 2016, 30–35.

Fischer, U., Berger, R. G., Hakansson, A., and Noble, A. C. (1996). The impact of dealcoholization on the flavour of wine - relating concentration of aroma compounds to

sensory data using PLS analysis. *Flavour Science. Recent Developments*. Royal Chemistry Society 197, 335–338.

Friedman, J., and Gore, J. (2017). Ecological systems biology: The dynamics of interacting populations. *Current Opinion in Systems Biology* 1, 114–121.

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., et al. (2016). Domestication and Divergence of Saccharomyces cerevisiae Beer Yeasts. *Cell* 166, 1397–1410.e16.

Gamero, A., Quintilla, R., Groenewald, M., Alkema, W., Boekhout, T., and Hazelwood, L. (2016). High-throughput screening of a large collection of non-conventional yeasts reveals their potential for aroma formation in food fermentation. *Food Microbiol.* 60, 147–159.

Ganucci, D., Guerrini, S., Mangani, S., Vincenzini, M., and Granchi, L. (2018). Quantifying the Effects of Ethanol and Temperature on the Fitness Advantage of Predominant Saccharomyces cerevisiae Strains Occurring in Spontaneous Wine Fermentations. *Front. Microbiol.* 9, 1563.

Geithe, C., Noe, F., Kreissl, J., and Krautwurst, D. (2017). The Broadly Tuned Odorant Receptor OR1A1 is Highly Selective for 3-Methyl-2,4-nonanedione, a Key Food Odorant in Aged Wines, Tea, and Other Foods. *Chem. Senses* 42, 181–193.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., et al. (1996). Life with 6000 genes. *Science* 274, 546, 563–7.

Goff, S. A., and Klee, H. J. (2006). Plant volatile compounds: sensory cues for health and nutritional value? *Science* 311, 815–819.

Goto-Yamamoto, N., Sawler, J., and Myles, S. (2015). Genetic Analysis of East Asian Grape Cultivars Suggests Hybridization with Wild Vitis. *PLoS One* 10, e0140841.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *Int. J. Hum. Comput. Stud.* 43, 907–928.

Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., et al. (2018). METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* 90, 3156–3164.

Hayasaka, Y., Dungey, K. A., Baldock, G. A., Kennison, K. R., and Wilkinson, K. L. (2010). Identification of a β-d-glucopyranoside precursor to guaiacol in grape juice following grapevine exposure to smoke. *Anal. Chim. Acta* 660, 143–148.

Henrici, A. T. (1933). Studies of Freshwater Bacteria: I. A Direct Microscopic Technique. *J. Bacteriol.* 25, 277–287.

Hernández-Orte, P., Ibarz, M. J., Cacho, J., and Ferreira, V. (2005). Effect of the addition of ammonium and amino acids to musts of Airen variety on aromatic composition and sensory properties of the obtained wine. *Food Chem.* 89, 163–174.

Hettich, R. L., Pan, C., Chourey, K., and Giannone, R. J. (2013). Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. *Anal. Chem.* 85, 4203–4214.

Hjelmeland, A. K., Zweigenbaum, J., and Ebeler, S. E. (2015). Profiling monoterpenol glycoconjugation in Vitis vinifera L. cv. Muscat of Alexandria using a novel putative compound database approach, high resolution mass spectrometry and collision induced dissociation fragmentation analysis. *Anal. Chim. Acta* 887, 138–147.

Hofmann, J., Hahm, H. S., Seeberger, P. H., and Pagel, K. (2015). Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature* 526, 241–244.

Holt, S., Mukherjee, V., Lievens, B., Verstrepen, K. J., and Thevelein, J. M. (2018). Bioflavoring by non-conventional yeasts in sequential beer fermentations. *Food Microbiol.* 72, 55–66.

Hranilovic, A., Gambetta, J. M., Schmidtke, L., Boss, P. K., Grbin, P. R., Masneuf-Pomarede, I., et al. (2018). Oenological traits of Lachancea thermotolerans show signs of domestication and allopatric differentiation. *Sci. Rep.* 8, 14812.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–93.

Kang, D. D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3, e1165.

Katarína, F., Katarína, M., Katarína, Ď., Ivan, Š., and Fedor, M. (2014). Influence of yeast strain on aromatic profile of Gewürztraminer wine. *LWT - Food Science and Technology* 59, 256–262.

Kennison, K. R., Wilkinson, K. L., Williams, H. G., Smith, J. H., and Gibberd, M. R. (2007). Smoke-derived Taint in Wine: Effect of Postharvest Smoke Exposure of Grapes on the Chemical Composition and Sensory Characteristics of Wine. *J. Agric. Food Chem.* 55, 10897–10901.

King, A., and Dickinson, J. R. (2000). Biotransformation of monoterpene alcohols by Saccharomyces cerevisiae, Torulaspora delbrueckii and Kluyveromyces lactis. *Yeast* 16, 499–506.

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018a). Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* 16, 410–422.

Knight, S. J., Klaere, S., Morrison-Whittle, P., and Goddard, M. R. (2018b). Fungal diversity during fermentation correlates with thiol concentration in wine: Fungal diversity correlates with wine thiol levels. *Aust. J. Grape Wine Res.* 24, 105–112.

Knight, S., Klaere, S., Fedrizzi, B., and Goddard, M. R. (2015). Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir. *Sci. Rep.* 5, 14233.

Kraus, J. K., Reed, G., and Villettaz, J.-C. (1984). Levures sèches actives de vinification. 2e partie et fin : utilisation et évaluation. *Conn. Vigne Vin.* 18:1–26.

Kurtzman, C., Fell, J. W., and Boekhout, T. (2011). *The Yeasts: A Taxonomic Study*. Elsevier.

Lawless, H. T., and Heymann, H. (2010). *Sensory Evaluation of Food: Principles and Practices*. Springer Science & Business Media.

Lee, S.-J., and Noble, A. C. (2003). Characterization of odor-active compounds in Californian chardonnay wines using GC-olfactometry and GC-mass spectrometry. *J. Agric. Food Chem.* 51, 8036–8044.

Lee, S.-J., and Noble, A. C. (2006). Use of Partial Least Squares Regression and Multidimensional Scaling on Aroma Models of California Chardonnay Wines. *Am. J. Enol. Vitic.* 57, 363–370.

Loscos, N., Hernandez-Orte, P., Cacho, J., and Ferreira, V. (2007). Release and formation of varietal aroma compounds during alcoholic fermentation from nonfloral grape odorless flavor precursors fractions. *J. Agric. Food Chem.* 55, 6674–6684.

Maggu, M., Winz, R., Kilmartin, P. A., Trought, M. C. T., and Nicolau, L. (2007). Effect of

skin contact and pressure on the composition of Sauvignon Blanc must. *J. Agric. Food Chem.* 55, 10281–10288.

Marais, J. (1983). Terpenes in the Aroma of Grapes and Wines: A Review. *S. Afr. J. Enol. Vitic.* 4, 49–58.

Masneuf-Pomarede, I., Bely, M., Marullo, P., and Albertin, W. (2015). The Genetics of Non-conventional Wine Yeasts: Current Knowledge and Future Challenges. *Front. Microbiol.* 6, 1563.

McGovern, P. E. (2013). *Ancient Wine: The Search for the Origins of Viniculture*. Princeton University Press.

Michel, M., Meier-Dörnberg, T., Jacob, F., Methner, F.-J., Wagner, R. S., and Hutzler, M. (2016). Review: Pure non- Saccharomyces starter cultures for beer fermentation with a focus on secondary metabolites and practical applications : Non-conventional yeast for beer fermentation. *J. Inst. Brew.* 122, 569–587.

Moio, L., and Etievant, P. X. (1995). Ethyl Anthranilate, Ethyl Cinnamate, 2,3-Dihydrocinnamate, and Methyl Anthranilate: Four Important Odorants Identified in Pinot noir Wines of Burgundy. *Am. J. Enol. Vitic.* 46, 392–398.

Morrison-Whittle, P., and Goddard, M. R. (2018). From vineyard to winery: a source map of microbial diversity driving wine fermentation. *Environ. Microbiol.* 20, 75–84.

Murat Eren, A., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319.

Nagel, C. W., Baranowski, J. D., Wulf, L. W., and Powers, J. R. (1979). The Hydroxycinnamic Acid Tartaric Acid Ester Content of Musts and Grape Varieties Grown in the Pacific Northwest. *Am. J. Enol. Vitic.* 30, 198–201.

Narayanasamy, S., Jarosz, Y., Muller, E. E. L., Heintz-Buschart, A., Herold, M., Kaysen, A., et al. (2016). IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. *Genome Biol.* 17, 260.

Nguyen, D. H., Nguyen, C. H., and Mamitsuka, H. (2018). Recent advances and prospects of computational methods for metabolite identification: a review with emphasis on machine learning approaches. *Brief. Bioinform.* doi:10.1093/bib/bby066.

Noble, A. C., and Ebeler, S. E. (2002). Use of multivariate statistics in understanding wine flavor. *Food Rev. Int.* 18, 1–20.

Noestheden, M., Dennis, E. G., Romero-Montalvo, E., DiLabio, G. A., and Zandberg, W. F. (2018). Detailed characterization of glycosylated sensory-active volatile phenols in smoke-exposed grapes and wine. *Food Chem.* 259, 147–156.

Nykänen, L. (1986). Formation and Occurrence of Flavor Compounds in Wine and Distilled Alcoholic Beverages. *Am. J. Enol. Vitic.* 37, 84–96.

Okamura, S., and Watanabe, M. (1981). Determination of Phenolic Cinnamates in White Wine and Their Effect on Wine Quality. *Agric. Biol. Chem.* 45, 2063–2070.

Pasteur, L. (1872). Nouvelles expériences pour démontrer que le germe de la levure qui fait le vin provient de l'extérieur des grains de raisin. *Comptes Rendus de l'Académie des Science de Paris* 75, 781–793.

Patti, G. J., Yanes, O., and Siuzdak, G. (2012). Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* 13, 263–269.

Peisl, B. Y. L., Schymanski, E. L., and Wilmes, P. (2018). Dark matter in host-microbiome metabolomics: Tackling the unknowns-A review. *Anal. Chim. Acta* 1037, 13–27.

Peters, K., Bradbury, J., Bergmann, S., Capuccini, M., Cascante, M., de Atauri, P., et al. (2018). PhenoMeNal: Processing and analysis of Metabolomics data in the Cloud. *Gigascience*. doi:10.1093/gigascience/giy149.

Poivet, E., Tahirova, N., Peterlin, Z., Xu, L., Zou, D.-J., Acree, T., et al. (2018). Functional odor classification through a medicinal chemistry approach. *Sci Adv* 4, eaao6086.

Portugal, C. B., de Silva, A. P., Bortoletto, A. M., and Alcarde, A. R. (2017). How native yeasts may influence the chemical profile of the Brazilian spirit, cachaça? *Food Res. Int.* 91, 18–25.

Pretorius, I. S. (2000). Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. *Yeast* 16, 675–729.

Rankine, B. C., and Lloyd, B. (1963). Quantitative assessment of dominance of added yeast in wine fermentations. *J. Sci. Food Agric.* 14, 793–798.

Rapp, A., and Mandery, H. (1986). Wine aroma. *Experientia* 42, 873–884.

Reed, G., and Nagodawithana, T. W. (1988). Technology of Yeast Usage in Winemaking. *Am. J. Enol. Vitic.* 39, 83–90.

Ribereau Gayon, P. (1999). Observations related to the causes and consequences of stuck fermentation in vinification [aeration of must, control of temperature, nutrient starvation]. *International Journal of Vine and Wine Sciences* 33, 39–48.

Robinson, J., and Harding, J. (2015). *The Oxford Companion to Wine*. Oxford University Press.

Roland, A., Schneider, R., Razungles, A., and Cavelier, F. (2011). Varietal thiols in wine: discovery, analysis and applications. *Chem. Rev.* 111, 7355–7376.

Román, S., Sánchez-Siles, L. M., and Siegrist, M. (2017). The importance of food naturalness for consumers: Results of a systematic review. *Trends Food Sci. Technol.* 67, 44–57.

Ruiz, J., Belda, I., Beisert, B., Navascués, E., Marquina, D., Calderón, F., et al. (2018). Analytical impact of Metschnikowia pulcherrima in the volatile profile of Verdejo white wines. *Appl. Microbiol. Biotechnol.* 102, 8501–8509.

Sáenz-Navajas, M.-P., Arias, I., Ferrero-Del-Teso, S., Fernández-Zurbano, P., Escudero, A., and Ferreira, V. (2018). Chemo-sensory approach for the identification of chemical compounds driving green character in red wines. *Food Res. Int.* 109, 138–148.

Sarry, J.-E., and Günata, Z. (2004). Plant and microbial glycoside hydrolases: Volatile release from glycosidic aroma precursors. *Food Chem.* 87, 509–521.

Schüttler, A. (2012). Influencing factors on aromatic typicality of wines from Vitis vinifera L. cv. *Riesling sensory, chemical and viticultural insights (Doctoral thesis). Université de Bordeaux and Justus Liebig University Giessen*.

Schüttler, A., Friedel, M., Jung, R., Rauhut, D., and Darriet, P. (2015). Characterizing aromatic typicality of Riesling wines: merging volatile compositional and sensory aspects. *Food Res. Int.* 69, 26–37.

Schymanski, E. L., Ruttkies, C., Krauss, M., Brouard, C., Kind, T., Dührkop, K., et al. (2017). Critical Assessment of Small Molecule Identification 2016: automated methods. *J. Cheminform.* 9, 22.

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., et al. (2017). Critical Assessment of Metagenome Interpretation-a benchmark of metagenomics software. *Nat. Methods* 14, 1063–1071.

Sefton, M. A., Skouroumounis, G. K., Elsey, G. M., and Taylor, D. K. (2011). Occurrence, sensory impact, formation, and fate of damascenone in grapes, wines, and other foods and beverages. *J. Agric. Food Chem.* 59, 9717–9746.

Serafim, F. A. T., Pereira-Filho, E. R., and Franco, D. W. (2016). Chemical data as markers of the geographical origins of sugarcane spirits. *Food Chem.* 196, 196–203.

Serafim, F. A. T., Seixas, F. R. F., Da Silva, A. A., Galinaro, C. A., Nascimento, E. S. P., Buchviser, S. F., et al. (2013). Correlation between chemical composition and sensory properties of Brazilian sugarcane spirits (cachaças). *J. Braz. Chem. Soc.* 24, 973–982.

Siebert, T. E., Barker, A., Pearson, W., Barter, S. R., de Barros Lopes, M. A., Darriet, P., et al. (2018). Volatile Compounds Related to 'Stone Fruit' Aroma Attributes in Viognier and Chardonnay Wines. *J. Agric. Food Chem.* 66, 2838–2850.

Simpson, R. F., and Miller, G. C. (1983). Aroma composition of aged Riesling wine. *Vitis* 22, 51–63.

Sirén, K., Fischer, U., and Vestner, J. (2019a). Automated supervised learning pipeline for non-targeted GC-MS data analysis. *Analytica Chimica Acta: X*, 100005.

Sirén, K., Mak, S. S. T., Fischer, U., Hansen, L. H., and Gilbert, M. T. P. (2019b). Multi-omics and potential applications in wine production. *Curr. Opin. Biotechnol.* 56, 172–178.

Skinkis, P. A., Bordelon, B. P., and Wood, K. V. (2008). Comparison of Monoterpene Constituents in Traminette, Gewürztraminer, and Riesling Winegrapes. *Am. J. Enol. Vitic.* 59, 440–445.

Stefanini, I., Albanese, D., Cavazza, A., Franciosi, E., De Filippo, C., Donati, C., et al. (2016). Dynamic changes in microbiota and mycobiota during spontaneous 'Vino Santo Trentino' fermentation. *Microb. Biotechnol.* 9, 195–208.

Stefanini, I., and Cavalieri, D. (2018). Metagenomic Approaches to Investigate the Contribution of the Vineyard Environment to the Quality of Wine Fermentation: Potentials and Difficulties. *Front. Microbiol.* 9, 991.

Sternes, P. R., Lee, D., Kutyna, D. R., and Borneman, A. R. (2017). A combined meta-barcoding and shotgun metagenomic analysis of spontaneous wine fermentation. *Gigascience* 6, 1–10.

Swiegers, J. H., Bartowsky, E. J., Henschke, P. A., and Pretorius, I. S. (2005). Yeast and bacterial modulation of wine aroma and flavour. *Aust. J. Grape Wine Res.* 11, 139–173.

Tempère, S., Marchal, A., Barbe, J.-C., Bely, M., Masneuf-Pomarede, I., Marullo, P., et al. (2018). The complexity of wine: clarifying the role of microorganisms. *Appl. Microbiol. Biotechnol.* 102, 3995–4007.

Tominaga, T., Baltenweck-Guyot, R., Des Gachons, C. P., and Dubourdieu, D. (2000). Contribution of Volatile Thiols to the Aromas of White Wines Made From Several Vitis vinifera Grape Varieties. *Am. J. Enol. Vitic.* 51, 178–181.

Tukey, J. W. (1980). We Need Both Exploratory and Confirmatory. *Am. Stat.* 34, 23–25.

van der Hooft, J. J. J., Wandy, J., Barrett, M. P., Burgess, K. E. V., and Rogers, S. (2016). Topic modeling for untargeted substructure exploration in metabolomics. *Proc. Natl. Acad. Sci. U. S. A.* 113, 13738–13743.

Varela, C. (2016). The impact of non-Saccharomyces yeasts in the production of alcoholic beverages. *Appl. Microbiol. Biotechnol.* 100, 9861–9874.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66–74.

Vèrette, E., Noble, A. C., and Somers, T. C. (1988). Hydroxycinnamates of Vitis vinifera: Sensory assessment in relation to bitterness in white wines. *J. Sci. Food Agric.* 45, 267–272.

Viant, M. R., Kurland, I. J., Jones, M. R., and Dunn, W. B. (2017). How close are we to complete annotation of metabolomes? *Curr. Opin. Chem. Biol.* 36, 64–69.

Vigentini, I., Maghradze, D., Petrozziello, M., Bonello, F., Mezzapelle, V., Valdetara, F., et al. (2016). Indigenous Georgian Wine-Associated Yeasts and Grape Cultivars to Edit the Wine Quality in a Precision Oenology Perspective. *Front. Microbiol.* 7, 352.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* 34, 828–837.

Williams, P. J., Strauss, C. R., Wilson, B., and Massy-Westropp, R. A. (1982). Use of C18 reversed-phase liquid chromatography for the isolation of monoterpene glycosides and nor-isoprenoid precursors from grape juice and wines. *J. Chromatogr. A* 235, 471–480.

Winterhalter, P., and Gök, R. (2013). 'TDN and β-Damascenone: Two Important Carotenoid Metabolites in Wine', in *ACS Symposium Series*, 125–137.

Winterhalter, P., Sefton, M. A., and Williams, P. J. (1990). Volatile C13-Norisoprenoid Compounds in Riesling Wine Are Generated From Multiple Precursors. *Am. J. Enol. Vitic.* 41, 277–283.

Winton, W., Ough, C. S., and Singleton, V. L. (1975). Relative Distinctiveness of Varietal Wines Estimated by the Ability of Trained Panelists to Name the Grape Variety Correctly. *Am. J. Enol. Vitic.* 26, 5–11.

Wolfender, J.-L., Nuzillard, J.-M., van der Hooft, J. J. J., Renault, J.-H., and Bertrand, S. (2018). Accelerating metabolite identification in natural product research: toward an ideal combination of LC-HRMS/MS and NMR profiling, in silico databases and chemometrics. *Anal. Chem.* doi:10.1021/acs.analchem.8b05112.

Wu, C., Liu, J., Zhao, P., Piringer, M., and Schauberger, G. (2016). Conversion of the chemical concentration of odorous mixtures into odour concentration and odour intensity: A comparison of methods. *Atmos. Environ.* 127, 283–292.

Ye, Y., and Tang, H. (2016). Utilizing de Bruijn graph of metagenome assembly for metatranscriptome analysis. *Bioinformatics* 32, 1001–1008.

Zepeda-Mendoza, M. L., Edwards, N. K., Madsen, M. G., Abel-Kistrup, M., Puetz, L., Sicheritz-Ponten, T., et al. (2018). Influence of Oenococcus oeni and Brettanomyces bruxellensis on Wine Microbial Taxonomic and Functional Potential Profiles. *Am. J. Enol. Vitic.* 69, 321–333.

Zheng, S.-J., Wang, Y.-L., Liu, P., Zhang, Z., Yu, L., Yuan, B.-F., et al. (2017). Stable isotope labeling-solid phase extraction-mass spectrometry analysis for profiling of thiols and aldehydes in beer. *Food Chem.* 237, 399–407.

Zoecklein, B. W., Marcy, J. E., Williams, J. M., and Jasinski, Y. (1997). Effect of Native Yeasts and Selected Strains of Saccharomyces cerevisiae on Glycosyl Glucose, Potential Volatile Terpenes, and Selected Aglycones of White Riesling (Vitis vinifera L.) Wines. *J. Food Compost. Anal.* 10, 55–65.

Chapter 1
Multi-omics and potential applications in wine production

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Multi-omics and potential applications in wine production

Kimmo Sirén[1,2], Sarah Siu Tze Mak[3], Ulrich Fischer[1], Lars Hestbjerg Hansen[4] and M Thomas P Gilbert[3,5]

The wine microbiome - that is the microbial communities associated with the fermentation of must, is one of the most important factors in transforming grapes to wine, including flavour and aroma. Recent developments in high throughput sequencing and other 'omics methodologies are rapidly changing the level and complexity of information that we are able to extract from the wine microbiome. This will significantly enhance not only our understanding of which microbes are present at the various stages of the grapevine growth and winemaking process, but also improve our understanding of the complex interactions between microbes, the substrate and environment, ultimately shaping wine production. In this perspective we describe the role and future potential of such techniques in wine production, and highlight the potential challenges that will be simultaneously faced.

Addresses
[1] Institute for Viticulture and Oenology, Dienstleistungszentrum Ländlicher Raum Rheinpfalz, Neustadt an der Weinstraße, Germany
[2] Department of Chemistry, University of Kaiserslautern, Kaiserslautern, Germany
[3] Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[4] Department of Environmental Science, Environmental Microbial Genomics Group, Aarhus University, Roskilde, Denmark
[5] Norwegian University of Science and Technology, University Museum, Trondheim, Norway

Corresponding author: Gilbert, M Thomas P (tgilbert@snm.ku.dk)

Current Opinion in Biotechnology 2019, **56**:172–178

This review comes from a themed issue on **Food biotechnology**

Edited by **Rute Neves** and **Herwig Bachmann**

https://doi.org/10.1016/j.copbio.2018.11.014

Introduction

Two microbes, — the yeast *Saccharomyces cerevisiae* and bacteria *Oenococcus oeni*, — have long been known to play a principal role alongside grape must in the creation of wine flavour and aroma, and as such are by far the two most well-studied microbes in winemaking. However, multiple other microbes are present during the winemaking process, whose potential relevance has been confirmed and even commercialized already, or are currently under investigation (Box 1). Microbes can impact the flavour and aroma of wine either in negative or positive ways, and these derive either from single microorganisms, or mixed consortia. Given that wine complexity is thought to increase with different microbial interactions [1], understanding, controlling and manipulating the microbial biodiversity present in winemaking is of crucial importance to the industry (Figure 1). Our aim here is to highlight the potential of different 'omics technologies for wine research, including recent examples of their application to winemaking, alongside others drawn from related fields for illustrative purposes. Although we acknowledge that the microbial influence in the soil and vineyard is also highly relevant, our main emphasis is on fermentation.

Types of multi-omics and their applications in wine production

Several 'omics technologies offer potential with regards to wine production (Table 1) as a complement to traditional culture-based techniques (Box 2). In the context of genetic data, High-throughput sequencing (HTS) tools represent two principal approaches. The first is amplicon-based sequencing (or so-called 'metabarcoding'), that focuses on specific components of the genomic information present. This is the modern manifestation of techniques such as Denaturant Gradient Gel Electrophoresis (DGGE) that first appeared in the early 1990s [2], and enables researchers to document the microbial community composition. The second is shotgun sequencing, that is increasingly the basis for true metagenomic and metatranscriptomic studies. Alongside these, several other multi-omics methods are starting to be considered, including metaproteomics and metabolomics, both of which assist in describing the functions of genes expressed by the microbial community, and their metabolic consequences. These methods allow higher throughput and new possibilities, as higher-resolution mass spectrometry (MS) combined with data-independent acquisition (DIA) approaches become more widely used [3••].

Metabarcoding

The heart of any metabarcoding study is PCR amplification of so-called 'barcoding' loci, using relatively generic primers, coupled to deep sequencing of the amplicons using

Box 1 Microbial mysteries in winemaking.

Which microbes underlie the winemaking process, and where in the process they originate from, has been of longstanding interest to winemaking research [42]. In principle, the microbes could originate from the vineyard, the winery, or from commercial starter cultures. A recent review about yeasts found in the vineyard and winery has concluded that factors shaping the microbial communities include at least the grapevine cultivar itself, the geographical location, climate, vineyard spraying regimes, technological practices, processing stage and season of the year [43]. The majority of such knowledge however, relates to isolated yeast cultures [44]. For example, it has been suggested that the capability of *S. cerevisiae* strains to adapt to alcohol and temperature might contribute to winery microbiota differentiation [45]. However, consensus about the contribution of commercial starters to cellar flora has not been reached yet [45–47]. Considerable future research is needed to investigate what role the broader microbial ecosystem plays, not least through integrating functional information on top of more basic taxonomic descriptions of which microbes are actually present and active.

an HTS platform such as the Illumina MiSeq. The first application to wine research targeted the bacterial 16S rRNA region [4], while subsequent studies have expanded the range of targets to include yeasts and other fungi, typically through targeting regions of the ITS or D1/D2 26S — something logical given the significant role of yeasts in winemaking, not least their influence on aroma [5].

Ultimately metabarcoding's power is to describe which microbes are present in any sample. Here, typical research questions include; where relevant microbes (whether beneficial or spoilage) enter the system, whether derived from the vineyard and grapes, fermentation or even bottling; and how their community profile changes through time. As an example, fungal communities vectored by *Drosophila* spp. have been found to differ among wineries, and thus the winery microbiota might be partially shaped or maintained by the fruit flies [6] - something that might extend the established notation of wine microbial regional fingerprints to winery flora [7–9]. We highlight that Belda *et al.* [10•], and Stefanini and Cavalier [11] have recently reviewed amplicon-based sequencing approaches with a particular focus on the applications in wine production, thus we do not discuss the findings in detail here. Instead we emphasize that despite metabarcoding's popularity, it has several shortcomings. Firstly, not all microbes PCR amplify equally well under any given primer system, thus possibly incurring systematic biases that exclude some taxa that may be important during wine fermentation. Secondly, taxonomic identification of the sequences generated is not straightforward [12]. While Edgar [13] presented a general evaluation of the accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences, one wine-specific example shows how metagenomic versus metabarcoding analyses indicate clear biases affecting the *Metschnikowia* genus with over representation in the amplification of ITS region [14•].

Metagenomics

In contrast, metagenomic approaches profile the whole genomic content of a target matrix [15,16••]. Thus this technique is particularly suited for questions that require insight into the overall gene content of the microbes - for example those attempting to predict microbial functional pathways [10•,17]. Thus, given the pangenome[6] of any species is bigger than any individual genome, it is challenging to assume certain genetic functions are present-based purely on knowing any microbial species is present. Similarly, the increased information gained enables better profiling of the microbial communities to species or even strain levels, and even verification of corresponding metabolic pathways [17]. Despite this promise, to date only two published studies exist in the wine production context (Table 1). One compares the use of metagenomics versus metabarcoding in assessing community structures of different spontaneous fermentation stages [14•], and the second documents the influence of different *Brettanomyces bruxellensis* strains and *O. oeni* starter cultures on total microbial communities during malolactic fermentation [18]. In general, metagenomic approaches promise to reveal the microbial community and microbial interaction in different stages of winemaking, similar to how they have been applied using metabarcoding (Table 1), but with the added benefit of yielding increased resolution, and in parallel, information into the functional pathways involved.

Metatranscriptomics

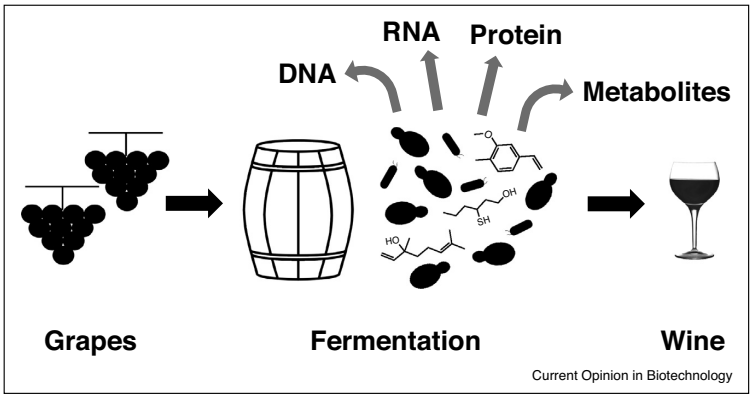
Metatranscriptomics refers to the measurement of total gene expression in a target matrix by extracting messenger RNA (mRNA) then converting it to cDNA using random hexamers or, in the case of Eukaryotes, poly-T primers that target the poly-A mRNA tail [18]. This data can shed light into the gene activity, thus functional relevance, of the target organisms within the matrix [10•,19–21]. Analyses can also be performed with stable isotope probing and the latest RNASeq in NGS platforms, for example to find specific targeted microbial transcriptomes in the samples [20] (Table 1). In general despite their potential to quantify the contribution of microbes to wine fermentation, research based on such methods is yet to be reported, although comparable examples have recently been published relating to the microbial activity during cheese ripening [22] and Chinese liquor production [23•].

Metaproteomics

Metaproteomics is the simultaneous study of multiple proteins expressed in any matrix, aiming to identify and quantify their presence, — information which simultaneously also improves the functional gene annotations and provides better insights into microbial protein-protein interactions inside the target matrix at any given time

[6] Pangenome refers to the full set of genes in all strains or clades in a species.

Figure 1



The 'omics of winemaking.

Table 1

Summary of multi-omics and their potential applications in wine production

Omics type	Target	Definition	Analytical technique(s)	Aim	Potential applications in wine production	References
Metagenomics	DNA	Profiling the whole genomic content of a target matrix applying untargeted shotgun sequencing approach	NGS/DNASeq	Gene contents, genetic pathways and identification of community	Changes in microbial community throughout wine production[a]; proactive spoilage detection; strain level detections; metabolomics impacts of whole microbial community	[14•,18]
Metatranscriptomics	Total RNA & mRNA	Measurement of total gene expressions in a target matrix by extracting total messenger RNA (mRNA)	NGS/ RNASeq; stable isotopes probing (SIP)	Gene expression and functions	Microbial pathway analyses in wine production; pro-active spoilage detection; microbial functions in wine production	
Metaproteomics	Protein	Identification and/or quantification of proteins expressed in a target matrix	Mass spectrometry can be coupled with chromatography	Functional activities	Haze formation; microbial protein monitoring during fermentations and infectious effect of unwanted microbes	
Metabolomics	Volatile and/or non-volatile compounds	Identification and/or quantification of metabolites using targeted and non-targeted analysis of chemical compounds in a target matrix	Mass spectrometry often coupled with chromatography; nuclear magnetic resonance (NMR)	Metabolic profiling and fingerprinting	Metabolite identification during fermentation[a]; characterization of metabolic pathways and precursors[a]; interaction among microbes; year and location classification studies[a]	[29–32]

[a] Indicates specific examples that have already been applied to wine production.

[24,25]. Although metaproteomics is currently underexploited in wine research, more traditional single protein characterising techniques have been applied to the most common protein related issues in wine production, for example the role of proteins in wine haze formation [26] and the impact of *Botrytis cinerea* proteins derived from infected grapes in initiating the detrimental gushing in finished sparkling wines [27]. Obvious applications include enhancing our understanding of microbial interactions through protein expression, and annotation of simultaneously generated nucleic acid data. Additionally, metaproteomics could be used to investigate taste-active peptides in wine, similar to the recently identified bitter peptides in cheese [28].

Box 2 Culture-based techniques.

Culture-based techniques are still standard in wine research. These are used for example, to study yeast interactions in inoculated wine fermentations as reviewed by Ciani *et al.* [44], or the dominance of *S. cerevisiae* during alcoholic fermentation as reviewed by Albergaria and Arneborg [48]. Others have used such methods better describe non-*Saccharomyces* yeasts that could be of potential interest in winemaking [49,50], but also to reveal previously undescribed yeasts that may have negative effects [51]. Traditionally such studies were based on initial isolation by culturing, followed by identification and enzymatic screening and characterization [52]. More recently whole genome sequencing (WGS) has expanded the toolbox with which to investigate yeasts and become a valuable tool for rapid characterization.

Given the power of HTS methods to describe genomes and associated transcriptomes, there is an obvious synergy between the two, especially when the aim is to genetically or even functionally profile the cultured microbes. For example, one recent study based on WGS of *S. cerevisiae* revealed how wine isolates in general, and commercial strains in particular, are highly similar at the genetic level [53]. This raises the question as to how they are nevertheless able to yield products with such variable profiles - or perhaps this fits with the claims from some authors that the use of commercial starter cultures has led to organoleptic standardization of wines [54]. A similar WGS-based approach on *O. oeni* however, has shown very different results, highlighting the existence of clear lineages among different fermented beverages [55], even with a distinctive regional variation [56].

Ultimately it is widely known that pure culture isolation and observation using artificial media often yields incomplete knowledge of a particular microbial flora [57]. Furthermore typically many of subdominant microbes that may add nuance to flavour cannot be cultivated [58]. Therefore, while in wine research culture-based methods have been successful in detecting and identifying single organisms, and these form the basis of studying their metabolic properties *in vitro*, the community is clearly biased towards culturable microorganisms [10•]. Hence there has been a strong interest in moving towards culture-independent methods, and their promise to radically alter how we view the microbial world [59].

Metabolomics

Metabolomic approaches aim to characterise multiple metabolites in a single matrix using mass spectrometry-based methods, enabling the identification and quantification of the underlying chemical compounds [20,29]. As with metaproteomics, metabolomic data can provide general proof of gene function, thus bolster the information gathered through metagenomic and transcriptomic surveys. Volatile and/or non-volatile constituents of the metabolic composition can be studied in either targeted and/or non-targeted fashion. Environmental factors and winemaking decisions have a strong impact on the microbial metabolic profiles; thus, metabolomics is useful in the investigation of dynamics between microbial communities and the matrix [20]. As outlined in Table 1, while major developments are happening in metabolomics, it has already been applied to wine production to study questions ranging from the cultivar differences, monitoring of the fermentation process and guiding of winemaking decision making, as well as the exploration of aroma and flavour variation by vintage [29–32].

Integrating multiple 'omics approaches

Different 'omics approaches reveal different molecular information of relevance and thereby enhance our understanding of the complex interactions between microbes, the substrate and environment. Given this, it stands to reason that the successful integration of different 'omics approaches into single studies will be particularly useful to wine production in the future. Although this approach has yet to be widely attempted in the context of wine making, there have been several recent studies of other in food fermentation systems where this approach has been able to enhance the understanding of overall microbial activities. Examples include the combining of metagenomics and metabolomic datasets to reveal the dominant flavour related microbes in different fermentation stages of Pu-Erh tea [33•] and during sausage ripening [34]; the combination of metagenomics and metatranscriptomic data to understand the activity of the microbial community during cheese ripening [35]; and the combination of metatranscriptomics and metabolomics to elucidate the functional changes in microbes under varies cheese ripening conditions [36]. With regards to winemaking, the sole current example that we are aware of is that reported by Bokulich *et al.* [8], who combined metabarcoding with metabolomic data while investigating the relationship between grape microbiome and regionality. Thus, overall the potential of combining these approaches remains underexploited. Although this is partly due to the previously prohibitively high cost of some of the methods, it is also because significant challenges remain to be addressed before any of the techniques become routine. One major challenge is not unique to wine data, and relates to the computational challenge of effectively dealing with the ever increasing resulting dataset sizes. A second is that the annotations of most microbial genes and genomes are not comprehensive [10•,37••] thus conferring considerable difficulty on those wishing to develop a picture of the microbial activities using metagenomic/metatranscriptomic analyses. Thirdly it is critical that studies are well-designed and appropriate statistical considerations are accounted for [38••]. The fourth challenge is much more basic, and relates to qualitative sample preparation — something that has been particularly limiting for metagenomic and metatransciptomic studies.

The challenge of sample preparation

The generation of reliable results during any 'omics application is heavily susceptible to both data quality in general, but also biases that may be linked to different preparation methods [10•]. Both grape must and wine are complex substrates, that contain large quantities of inhibitors such as polyphenols and polysaccharides [10•,39]. Unfortunately while commercial DNA extraction kits have been designed to solve similar difficulties in other substrates, there is no such optimised solution for wine, and instead more customised extraction methods are needed, including extra purifications to remove the inhibitors [10•,40]. Regrettably, however, this can result in very low final DNA/RNA yields,

which in turn risks heavily biasing the recovered genetic profiles. It is not uncommon that researchers need to undertake specific nucleic acid extraction comparisons prior to commencing larger projects. Another challenge, for at least metagenomic studies, is the co-isolation of large amounts of non-microbial genetic material derived from the grape vine itself. Although once sequenced, this data can be removed bioinformatically during subsequent analysis, this naturally reduces overall efficiency, mandating elevated sequencing depth, thus overall cost [10•]. Additionally, when sequencing microbial transcriptomes, because rRNA sequences are so abundant, the actual mRNA may only represent ~5% of the total RNA extracted. Because Prokaryotic mRNA lacks poly-A tails, they cannot be enriched using methods applied to Eukaryotes [19]. Thus, specific mRNA enrichment, or microbial rRNA and host mRNA depletion methods are needed (all of which are commercially available, although perhaps not equally efficient across microbial species [41]), which also selectively bias the results [20]. We highlight however, that neither approaches have been widely tested with regards to the wine transcriptome, hence tests will be needed to guide future studies on their suitability for resolving the problem.

Conclusions

Ultimately, as wine is the product of complex microbial communities, studies that focus solely on single isolated strains are not sufficient for fully understanding its formation. While still only in their infancy, it is clear that in the future the application of multi-omics approaches will be critical for improving our understanding of the wine microbiome. In the future, multi-omics could be used for example to reveal effects upon specific stress conditions or to identify biological pathways leading to specific aroma formations. Such approaches are of great interest not only for wine research but also for the whole wine community and other food industries where complex microbial networks have an essential impact.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors would like to thank the Horizon 2020 Programme of the European Commission within the Marie Skłodowska-Curie Innovative Training Network 'MicroWine' (grant number 643063) for financial support, and the creation of barrel graphics by Alex Muravev and Carlos Dias from the Noun Project.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Tempère S, Marchal A, Barbe J-C, Bely M, Masneuf-Pomarede I, Marullo P, Albertin W: **The complexity of wine: clarifying the role of microorganisms**. *Appl Microbiol Biotechnol* 2018, **102**: 3995-4007.

2. Muyzer G, de Waal EC, Uitterlinden AG: **Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA**. *Appl Environ Microbiol* 1993, **59**:695-700.

3. •• Bonner R, Hopfgartner G: **SWATH data independent acquisition mass spectrometry for metabolomics**. *Trends Analyt Chem* 2018 http://dx.doi.org/10.1016/j.trac.2018.10.014.
Review describes the SWATH-MS technique and its' potential applications to small molecules analyses.

4. Bokulich NA, Joseph CML, Allen G, Benson AK, Mills DA: **Next-generation sequencing reveals significant bacterial diversity of botrytized wine**. *PLoS One* 2012, **7**:e36357.

5. Gammacurta M, Marchand S, Moine V, de Revel G: **Influence of different yeast/lactic acid bacteria combinations on the aromatic profile of red Bordeaux wine**. *J Sci Food Agric* 2017, **97**:4046-4057.

6. Quan AS, Eisen MB: **The ecology of the *Drosophila*-yeast mutualism in wineries**. *PLoS One* 2018, **13**:e0196440.

7. Knight S, Klaere S, Fedrizzi B, Goddard MR: **Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir**. *Sci Rep* 2015, **5**:14233.

8. Bokulich NA, Collins TS, Masarweh C, Allen G, Heymann H, Ebeler SE, Mills DA: **Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics**. *MBio* 2016, **7** e00631-16.

9. Morrison-Whittle P, Goddard MR: **From vineyard to winery: a source map of microbial diversity driving wine fermentation**. *Environ Microbiol* 2018, **20**:75-84.

10. • Belda I, Zarraonaindia I, Perisin M, Palacios A, Acedo A: **From vineyard soil to wine fermentation: microbiome approximations to explain the "terroir" concept**. *Front Microbiol* 2017, **8**:821.
Review focuses on NGS technologies used in grapevine and wine studies.

11. Stefanini I, Cavalieri D: **Metagenomic approaches to investigate the contribution of the vineyard environment to the quality of wine fermentation: potentials and difficulties**. *Front Microbiol* 2018, **9**:991.

12. Colabella C, Corte L, Roscini L, Bassetti M, Tascini C, Mellor JC, Meyer W, Robert V, Vu D, Cardinali G: **NGS barcode sequencing in taxonomy and diagnostics, an application in "*Candida*" pathogenic yeasts with a metagenomic perspective**. *IMA Fungus* 2018, **9**:91-105.

13. Edgar RC: **Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences**. *PeerJ* 2018, **6**:e4652.

14. • Sternes PR, Lee D, Kutyna DR, Borneman AR: **A combined meta-barcoding and shotgun metagenomic analysis of spontaneous wine fermentation**. *Gigascience* 2017, **6**:1-10.
First comparison of metagenomic and metabarcoding approaches in wine.

15. Ferrocino I, Cocolin L: **Current perspectives in food-based studies exploiting multi-omics approaches**. *Curr Opin Food Sci* 2017, **13**:10-15.

16. •• Mardanov AV, Kadnikov VV, Ravin NV: **Chapter 1 - Metagenomics: a paradigm shift in microbiology**. In *Metagenomics*. Edited by Nagarajan M. Academic Press; 2018: 1-13.
Comprehensive overview of metagenomics history.

17. Walsh AM, Crispie F, Claesson MJ, Cotter PD: **Translating omics to food microbiology**. *Annu Rev Food Sci Technol* 2017, **8**: 113-134.

18. Zepeda-Mendoza ML, Edwards NK, Madsen MG, Abel-Kistrup M, Puetz L, Sicheritz-Ponten T, Swiegers JH: **Influence of *Oenococcus oeni* and *Brettanomyces bruxellensis* on wine microbial taxonomic and functional potential profile**. *Am J Enol Vitic* 2018, **69**:321-333.

19. Bashiardes S, Zilberman-Schapira G, Elinav E: **Use of metatranscriptomics in microbiome research**. *Bioinform Biol Insights* 2016, **10**:19-25.

20. Aguiar-Pulido V, Huang W, Suarez-Ulloa V, Cickovski T, Mathee K, Narasimhan G: **Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis**. *Evol Bioinform Online* 2016, **12**:5-16.

21. De Filippis F, Parente E, Ercolini D: **Recent past, present, and future of the food microbiome**. *Annu Rev Food Sci Technol* 2018, **9**:589-608.

22. Yeluri Jonnala BR, McSweeney PLH, Sheehan JJ, Cotter PD: **Sequencing of the cheese microbiome and its relevance to industry**. *Front Microbiol* 2018, **9**:1020.

23. • Song Z, Du H, Zhang Y, Xu Y: **Unraveling core functional microbiota in traditional solid-state fermentation by high-throughput amplicons and metatranscriptomics sequencing**. *Front Microbiol* 2017, **8**:1294.
Integrating metabarcoding data with metatranscriptomics data in Chinese liquor.

24. Zhao M, Zhang D-L, Su X-Q, Duan S-M, Wan J-Q, Yuan W-X, Liu B-Y, Ma Y, Pan Y-H: **An integrated metagenomics/metaproteomics investigation of the microbial communities and enzymes in solid-state fermentation of Pu-erh tea**. *Sci Rep* 2015, **5**:10117.

25. Liu Y, Rousseaux S, Tourdot-Maréchal R, Sadoudi M, Gougeon R, Schmitt-Kopplin P, Alexandre H: **Wine microbiome: a dynamic world of microbial interactions**. *Crit Rev Food Sci Nutr* 2017, **57**:856-873.

26. Ferreira RB, Piçarra-Pereira MA, Monteiro S, Loureiro VB, Teixeira AR: **The wine proteins**. *Trends Food Sci Technol* 2001, **12**:230-239.

27. Kupfer VM, Vogt EI, Ziegler T, Vogel RF, Niessen L: **Comparative protein profile analysis of wines made from *Botrytis cinerea* infected and healthy grapes reveals a novel biomarker for gushing in sparkling wine**. *Food Res Int* 2017, **99**:501-509.

28. Sebald K, Dunkel A, Schäfer J, Hinrichs J, Hofmann T: **Sensoproteomics: a new approach for the identification of taste-active peptides in fermented foods**. *J Agric Food Chem* 2018, **66**:11092-11104.

29. Cozzolino D: **Metabolomics in grape and wine: definition, current status and future prospects**. *Food Anal Methods* 2016, **9**:2986-2997.

30. Alañón ME, Pérez-Coello MS, Marina ML: **Wine science in the metabolomics era**. *Trends Analyt Chem* 2015, **74**:1-20.

31. Arapitsas P, Guella G, Mattivi F: **The impact of SO2 on wine flavanols and indoles in relation to wine style and age**. *Sci Rep* 2018, **8**:858.

32. Ma Y, Tanaka N, Vaniya A, Kind T, Fiehn O: **Ultrafast polyphenol metabolomics of red wines using MicroLC-MS/MS**. *J Agric Food Chem* 2016, **64**:505-512.

33. • Li Z, Feng C, Luo X, Yao H, Zhang D, Zhang T: **Revealing the influence of microbiota on the quality of Pu-erh tea during fermentation process by shotgun metagenomic and metabolomic analysis**. *Food Microbiol* 2018, **76**:405-415.
First integration of metagenomic and metabolomic approaches in fermented beverage.

34. Ferrocino I, Bellio A, Giordano M, Macori G, Romano A, Rantsiou K, Decastelli L, Cocolin L: **Shotgun metagenomics and volatilome profile of the microbiota of fermented sausages**. *Appl Environ Microbiol* 2018, **84** e02120-17.

35. Dugat-Bony E, Straub C, Teissandier A, Onésime D, Loux V, Monnet C, Irlinger F, Landaud S, Leclercq-Perlat M-N, Bento P *et al.*: **Overview of a surface-ripened cheese community functioning by meta-omics analyses**. *PLoS One* 2015, **10**:e0124360.

36. De Filippis F, Genovese A, Ferranti P, Gilbert JA, Ercolini D: **Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate**. *Sci Rep* 2016, **6**:21871.

37. •• Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S, Dröge J, Gregor I, Majda S, Fiedler J, Dahms E *et al.*: **Critical assessment of metagenome interpretation—a benchmark of metagenomics software**. *Nat Methods* 2017, **14**:1063.
Benchmarking challenge for metagenomic software.

38. •• Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, Gonzalez A, Kosciolek T, McCall L-I, McDonald D *et al.*: **Best practices for analysing microbiomes**. *Nat Rev Microbiol* 2018, **16**:410-422.
Outstanding overview of best practices for performing a microbiome study.

39. Işçi B, Kalkan Yildirim H, Altindisli A: **Evaluation of methods for DNA extraction from must and wine**. *J Inst Brew* 2014, **120**:238-243.

40. Lever MA, Torti A, Eickenbusch P, Michaud AB, antl-Temkiv T, Jørgensen BB: **A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types**. *Front Microbiol* 2015, **6**:476.

41. Herbert ZT, Kershner JP, Butty VL, Thimmapuram J, Choudhari S, Alekseyev YO, Fan J, Podnar JW, Wilcox E, Gipson J *et al.*: **Cross-site comparison of ribosomal depletion kits for Illumina RNAseq library construction**. *BMC Genomics* 2018, **19**:199.

42. Pasteur L: **Nouvelles expériences pour démontrer que le germe de la levure qui fait le vin provient de l'extérieur des grains de raisin**. *CR Searces Acad Sci Paris* 1872, **75**:781-793.

43. Varela C, Borneman AR: **Yeasts found in vineyards and wineries**. *Yeast* 2017, **34**:111-128.

44. Ciani M, Capece A, Comitini F, Canonico L, Siesto G, Romano P: **Yeast interactions in inoculated wine fermentation**. *Front Microbiol* 2016, **7**:555.

45. Ganucci D, Guerrini S, Mangani S, Vincenzini M, Granchi L: **Quantifying the effects of ethanol and temperature on the fitness advantage of predominant *Saccharomyces cerevisiae* strains occurring in spontaneous wine fermentations**. *Front Microbiol* 2018, **9**:1563.

46. Börlin M, Venet P, Claisse O, Salin F, Legras J-L, Masneuf-Pomarede I: **Cellar-associated *Saccharomyces cerevisiae* population structure revealed high-level diversity and perennial persistence at sauternes wine estates**. *Appl Environ Microbiol* 2016, **82**:2909-2918.

47. Martiniuk JT, Pacheco B, Russell G, Tong S, Backstrom I, Measday V: **Impact of commercial strain use on *Saccharomyces cerevisiae* population structure and dynamics in pinot noir vineyards and spontaneous fermentations of a canadian winery**. *PLoS One* 2016, **11**:e0160259.

48. Albergaria H, Arneborg N: **Dominance of *Saccharomyces cerevisiae* in alcoholic fermentation processes: role of physiological fitness and microbial interactions**. *Appl Microbiol Biotechnol* 2016, **100**:2035-2046.

49. Ruiz J, Belda I, Beisert B, Navascués E, Marquina D, Calderón F, Rauhut D, Santos A, Benito S: **Analytical impact of *Metschnikowia pulcherrima* in the volatile profile of Verdejo white wines**. *Appl Microbiol Biotechnol* 2018, **102**:8501-8509.

50. Jood I, Hoff JW, Setati ME: **Evaluating fermentation characteristics of *Kazachstania* spp. and their potential influence on wine quality**. *World J Microbiol Biotechnol* 2017, **33**:129.

51. Prendes LP, Merín MG, Fontana AR, Bottini RA, Ramirez ML, Morata de Ambrosini VI: **Isolation, identification and selection of antagonistic yeast against *Alternaria alternata* infection and tenuazonic acid production in wine grapes from Argentina**. *Int J Food Microbiol* 2018, **266**:14-20.

52. Belda I, Ruiz J, Alastruey-Izquierdo A, Navascués E, Marquina D, Santos A: **Unraveling the enzymatic basis of wine "Flavorome": a phylo-functional study of wine related yeast species**. *Front Microbiol* 2016, **7**:243.

53. Borneman AR, Forgan AH, Kolouchova R, Fraser JA, Schmidt SA: **Whole genome comparison reveals high levels of inbreeding and strain redundancy across the spectrum of commercial wine strains of *Saccharomyces cerevisiae***. *G3* 2016, **6**:957-971.

54. Vigentini I, Maghradze D, Petrozziello M, Bonello F, Mezzapelle V, Valdetara F, Failla O, Foschino R: **Indigenous georgian wine-associated yeasts and grape cultivars to edit the wine quality in a precision oenology perspective**. *Front Microbiol* 2016, **7**:352.

55. Campbell-Sills H, El Khoury M, Favier M, Romano A, Biasioli F, Spano G, Sherman DJ, Bouchez O, Coton E, Coton M *et al.*: **Phylogenomic analysis of *Oenococcus oeni* reveals specific domestication of strains to cider and wines**. *Genome Biol Evol* 2015, **7**:1506-1518.

56. El Khoury M, Campbell-Sills H, Salin F, Guichoux E, Claisse O, Lucas PM: **Biogeography of *Oenococcus oeni* reveals distinctive but nonspecific populations in wine-producing regions**. *Appl Environ Microbiol* 2017, **83** e02322–16.

57. Henrici AT: **Studies of freshwater bacteria: I. A direct microscopic technique**. *J Bacteriol* 1933, **25**:277-287.

58. Amann RI, Ludwig W, Schleifer KH: **Phylogenetic identification and in situ detection of individual microbial cells without cultivation**. *Microbiol Rev* 1995, **59**:143-169.

59. Lleixà J, Kioroglou D, Mas A, Portillo MDC: **Microbiome dynamics during spontaneous fermentations of sound grapes in comparison with sour rot and *Botrytis* infected grapes**. *Int J Food Microbiol* 2018, **281**:36-46.

Chapter 2

Automated supervised learning pipeline for non-targeted GC-MS data analysis

Analytica Chimica Acta: X 1 (2019) 100005

Contents lists available at ScienceDirect

Analytica Chimica Acta: X

journal homepage: www.journals.elsevier.com/analytica-chimica-acta-x

Automated supervised learning pipeline for non-targeted GC-MS data analysis

Kimmo Sirén [a, b], Ulrich Fischer [a], Jochen Vestner [a, *]

[a] Institute for Viticulture and Oenology, DLR Rheinpfalz, Breitenweg 71, D-67435, Neustadt, Germany
[b] Department of Chemistry, University of Kaiserslautern, Erwin-Schroedinger-Strasse 52, D-67663, Kaiserslautern, Germany

ARTICLE INFO

Article history:
Received 15 November 2018
Received in revised form
21 December 2018
Accepted 2 January 2019
Available online 10 January 2019

Keywords:
Metabolomics
Chemometrics
Tensor decomposition
Machine learning
Classification
Exploratory data analysis

ABSTRACT

Non-targeted analysis is nowadays applied in many different domains of analytical chemistry such as metabolomics, environmental and food analysis. Conventional processing strategies for GC-MS data include baseline correction, feature detection, and retention time alignment before multivariate modeling. These techniques can be prone to errors and therefore time-consuming manual corrections are generally necessary. We introduce here a novel fully automated approach to non-targeted GC-MS data processing. This new approach avoids feature extraction and retention time alignment. Supervised machine learning on decomposed tensors of segmented chromatographic raw data signal is used to rank regions in the chromatograms contributing to differentiation between sample classes. The performance of this novel data analysis approach is demonstrated on three published datasets.

1. Introduction

High-throughput approaches combined with non-targeted analysis are increasingly used in several different research disciplines such as food and environmental sciences, systems biology and metabolomics. In contrast to targeted analysis methods, the non-targeted approaches are inherently more holistic and therefore provide more comprehensive overview of the sample composition. These approaches are more hypothesis generating as they do not rely on *a priori* defined set of compounds but take known and unknown compounds into account.

The main objective of non-targeted studies is to discover molecules, which distinguish between groups or classes of samples, such as biomarkers from large sparse datasets [1–3]. As high-throughput approaches result in large sample sizes, this objective can be seen as a machine learning classification problem. For LC-MS data open source software such as XCMS [4], MS-DIAL [5] and many others are usually applied [6]. GC-MS data analysis commonly relies on vendor software which is often not applicable or compatible to data formats of other vendors [7]. The freely available software tools used in metabolomics have been recently listed [6]. Due to the popularity of XCMS for LC-MS, many have applied it for GC-MS, though however, the manual parameter optimisation remains vague [8]. Common data analysis approaches for non-targeted GC-MS data use feature detection in single ion chromatograms of individual samples in order to extract quantitative information resulting in concatenated data frames such as peak tables [1]. Typically downstream approaches such as multivariate analysis techniques e.g. unsupervised principal component analysis (PCA) and supervised partial least squares discriminant analysis (PLS-DA) are often subsequently used to analyze and visualize these feature tables. Automated feature detection however remains troublesome due to coelution, retention time shifts and potential erroneous peak picking and/or peak assignment. Relevant information can get lost by using intensity thresholds to differentiate between analytical signals and noise. Moreover, feature detection and alignment of chromatograms are difficult to automate as algorithm settings have to be optimized and validated resulting in more hands-on time. This also reduces reproducibility of the studies.

Several alternative approaches to conventional non-targeted GC-MS data analysis have been developed [9–23]. These approaches aim at better extraction of information and underlying patterns by using the GC-MS raw data signals (retention time × mass-to-charge ratio) as chromatographic fingerprints for modeling. In this way, feature detection is circumvented. Some of

* Corresponding author.
E-mail address: jochen.vestner@dlr.rlp.de (J. Vestner).

https://doi.org/10.1016/j.acax.2019.100005

Abbreviations

PARAFAC	Parallel Factor Analysis
PLS-DA	Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis
SSCP	Sum of Squares and Cross Product
t-SNE	t-Distributed Stochastic Neighbor Embedding
TIC	Total Ion Chromatogram
RMS-noise	Root Mean Square Noise consensus score

these alternative approaches are avoiding retention time alignment by implementing transformation of raw data signal to particular distance matrices [9,10,13,16–19]. A recent clustering approach [22] to extract features from single samples has been updated to fit MZMine 2 [24] workflow. Additionally, other recent full workflow approaches rely on baseline and time shift corrections or other preprocessing methods [21,25–27]. Multiway approaches to GC-MS data [9,16,18,20] such as PARAFAC has been evaluated against multivariate curve resolution by alternating least squares (MCR-ALS) which are the two most well-known mixture resolving methods. [28]. Recent approaches using alternative segmentation of the GC-MS raw data signal and subsequent feature se lection have been proposed [20,23] These approaches rely on normalized data [23] or focus on single samples [20]. However, many recent approaches that use multivariate statistics or machine learning still continue to rely on preprocessed peak tables instead of using the raw signals for modeling [29–34]. Vestner et al. [10] based their data analysis approach on avoiding classical feature extraction and alignment by segmenting chromatograms along the retention time axis and transforming the two dimensional chromatographic segments to sums of squares and cross-product (SSCP) matrices of the mass channels. In this study we expand this idea further by using supervised learning methods to employ the neglected power of sample sizes and *a priori* knowledge of group memberships on SSCP matrices of automatically chosen segments. We minimize hands-on time by focusing on ranking segments in the chromatogram. Ultimately, the ranking is established on the ability of individual segments to differentiate between classes. Downstream analysis of segments contributing to class differentiation is similar to conventional approaches and can include visual exploration or more sophisticated deconvolution of mass spectra of the relevant regions can be performed if needed [19,35]. Moreover, chemical compounds in the important segments can subsequently be tentatively identified based on retention indices and (resolved) mass spectra.

The main objective of this research was the development of a fast and fully automated approach avoiding time consuming optimization and parameter tuning. The performance of this novel supervised learning approach is demonstrated on three published datasets by comparing the approach to the published results.

2. Material and methods

2.1. Theory

An overview of the workflow is presented in Fig. 1. The workflow is based on following ideas: Segmentation of the whole dataset, construction of a classifier model for the transformed raw data of each segment separately, and the ranking of segments according to their importance to discriminate between classes of samples using model performance metrics. Retention time alignment is circumvented by an automated segmentation of chromatograms along the retention time axis. Subsequent modeling of the full GC-MS raw

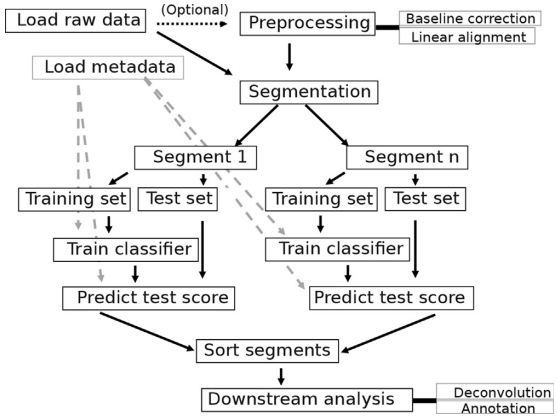


Fig. 1. Flowchart of the non-targeted GC-MS data analysis workflow. Gray arrows mark stages where metadata is needed. The whole analysis is automatic and does not require manual work steps or parameter tuning.

data signals of each individual segment makes feature extraction unnecessary. For each segment, chromatographic raw data signals of all samples are transformed into SSCP matrices following the approach described by Vestner et al. [10]. These SSCP matrices for all samples are merged into a tensor of order three, and decomposed using tensor decomposition. Next, a supervised machine learning pipeline based on gradient boosted tree classification is used on the decomposed tensor. The performance of prediction models are evaluated in order to establish the ranking of the segments. Classification metrics provide information on the importance of each segment on the differentiation among sample classes.

In detail, automated segmentation of chromatograms is achieved by finding local minima of summed total ion chromatograms (TICs) of all samples. With regular injection precision of GC-MS instrumentation, baseline correction and retention time alignment are generally not needed. Larger retention time shifts in GC-MS analysis indicate instability of the system and should rather be solved by optimizing the instrument performance instead of correcting data. However, linear alignment of chromatograms can be used to minimize inter-segment shifting of peaks. Whereas intra-segment shifting is dealt with by transforming the raw data signal matrix (retention time $\times$ mass-to-charge ratio) of each segment and sample to SSCP matrices. For each segment, the SSCP matrices of all samples are stacked together into tensors of order three. Each of these tensors is then individually decomposed using tucker decomposition [36]. Ultimately, the tensor decomposition sample loadings can be interpreted as a representation of the major source of variation among samples in each segment. The tensor decomposition core sizes are optimized to retain 99% variation by selecting enough ranks. The sample dimension loadings are then used to solve the supervised learning classification problem. Thus, a supervised classification model (classifier) is trained based on predefined classes for each segment individually resulting in as many models as segments. Principally, any classifier could be used but tree ensemble models have been found to retain higher order interactions between features and no inter-sample scaling of inputs is required [37].

Finally, the ranking of the segments is attained by evaluating the performance of each model (segment) using custom classification metrics. This ranking of the models according to their importance reflects the amount of information relevant in each segment - thus allowing extraction of the most relevant segments. We follow

established definitions for classification metrics in this paper [38]. The metrics used here are chosen to be applicable to evaluate both binary and multi-class classification models. Classical metrics such as precision, recall and F-score are generally regarded to be insufficient to estimate the performance of classification [39]. Requirements for metrics are as follows: First, individual classes should be highly recoverable. Second, model stability needs to be evaluated. A similar approach has been suggested to classify transcriptomics data using Precision-Recall (PR) curves and Receiver Operating Characteristics (ROC) area under curves (AUC) [40]. PR curves present more informative picture of model performance even though PR and ROC spaces are deeply connected. Therefore, if a curve dominates in ROC space it also dominates in PR space [41]. In order to control class imbalances, micro-average scores are used, so that by the aggregation of all class contributions the average is calculated. The multi-class ROC and AUC calculations were based on Provost and Domingos [42] approach as described by Fawcett [38]. Micro average of AUC and the maximum average value of the PR curve area are used in two dimensional fashion as described by Carbonero-Ruz et al. [43].

After ranking, the most important segments discriminating between sample classes can be further downstream analyzed to extract more chemically meaningful information. Moreover, Principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE) [44] or clustered heat maps based on similarity and dissimilarity matrices can be used to visually investigate the relations between segments and samples. Additionally, t-SNE can be useful in order to retain both global and local structure which is not possible with traditional dimensionality reduction techniques such as PCA [44]. Subsequently, further downstream processing could include deconvolution of mass spectra using e.g. PARAFAC2 as recently suggested by Tian et al. [45] followed by metabolite annotation process.

2.2. Methods

The here described approach is entirely written in Python 3 programming language [46]. GC-MS files were imported as NetCDF-format [47]. Neither baseline correction, nor linear alignment were applied. For the segmentation, a consensus chromatogram is obtained by summing the TICs of all samples and a consensus score of the root mean square noise (RMS-noise) is estimated. More specifically, the RMS-noise consensus score is estimated by calculating the square root of the summed intensities of each sample and taking the mean value of all the samples scaled to a percentage. Local minima are found by comparing the potential candidate point to the following data points while taking into account the RMS-noise. Note that the lower the RMS-noise value, the more minima are found. A modified script from https://gist.github.com/sixtenbe/1178136 was applied.

In every segment for each sample, a SSCP matrix was calculated. Subsequently, the SSCP matrices of each sample were concatenated to a tensor of order 3 for each segment. These tensors were individually decomposed using tucker decomposition via higher order orthogonal iteration [48,49]. Sample loadings were used to build a supervised learning classification problem. In addition, for model validation, leave-one-out cross-validation was chosen due to the relative low amount of samples in one of the dataset (Table 1).

For the statistical learning pipeline, a tree booster classifier XGBOOST was employed [37]. The default model parameters of the algorithm were retained without further fine-tuning to showcase the approach. The following parameters were used: learning_rate = 0.1, max_depth = 3, min_child_weight = 1, n_estimators = 100, reg_alpha = 0, reg_lambda = 1, scale_pos_weight = 1. Probabilistic function softmax was used as an objective function. Evaluation of the model cycles was set with multiple logistic loss criteria for multiple classes as in logistic regression for the training loss.

Table 1
Description of the datasets used.

Dataset	Samples	Number of classes	Project	Reference
Wine	39	3	Not publicly available	[10]
Urea	160	4	MTBLS71	[55]
Rice	79	20	MTBLS288	[56]

In order to rank and choose segments for further downstream analysis, ROC, AUC and PR curves were calculated from the true and the predicted probabilities of classes for every segment [50]. Segments with micro-average of ROC curve threshold and maximum class PR curve value higher than 0.9 were collected and further evaluated. For downstream examples the learning pipeline was rerun keeping only these selected segments. The parameters for t-SNE were set as follows: perplexity 10, early exaggeration 2, and learning rate 10. Acceleration of the computation was done using Barnes-Hut approach and initializing with dimensionality reduction with PCA [51]. The multidimensional data was visualized using Matplotlib and Seaborn plotting libraries [52,53].

2.3. Application

The fully automatized non-targeted GC-MS data analysis approach reported herein was tested with three different and independent, already published datasets. Two data sets are publicly available through the Metabolights database [54]. Using published data for the verification of the approach has two main advantages. Firstly, the outcome of the data analysis can be directly compared with already published findings. Secondly, multiple datasets from different domains of analytical chemistry can be used for confirming the applicability of the new approach in these domains.

For an overview of the datasets see Table 1. The first dataset consists of GC-MS fingerprinting analysis of volatile constituents of Cabernet Sauvignon wines submitted to different fermentation scenarios. The second dataset derives from a metabolomics study on the impact of urease treatment of urine samples prior to GC-MS [55] analysis. The third dataset originates from a metabolomics study on rice focusing on metabolic shifts during rice grain development [56]. Our workflow, shown in Fig. 1, was applied to all datasets in the same way. The only interactions between the user and the software are: providing the target classes and optionally controlling the segmentation with desired level of RMS-noise and how many data points are included in the segmentation algorithm. The workflow is therefore completely free of user interaction during the data analysis. For the workflow comparison the segmentation parameters were kept default as following: data points: 15; RMS-noise: 1.

2.4. Impact of segmentation on algorithm performance

In the following, the accuracy of finding relevant segment is assessed. In this regards, the impact of the number of segments on the performance of the approach was investigated for all three datasets without linear alignment as follows: Segment sizes were varied for the automated segmentation by changing the amount of RMS-noise (RMS-noise ε {100, 10, 1, 0.1}). All segments with micro-average AUC score higher than 0.9 were selected. Additionally, cophenetic correlation coefficients were calculated to estimate the quality of the hierarchical clustering dendrogram by preserving the pairwise correlation distances with average clustering method

between the original data [57,58].

For the urea dataset the same approach was followed with a robust linear retention time alignment using openCV module [59]. Ultimately, the same dataset was also further downstream analyzed by deconvolution of the important segments using PARAFAC2 algorithm of the multiway package [60] in R (version 3.4.1) [61]. Random forest based [62] feature selection wrapper algorithm [63] was used to retain most relevant deconvoluted peaks contributing to the separation between the four classes.

3. Results and discussion

A comparison of the outcome of our new approach to published results for the three data sets is presented in the following. The major focus of our new algorithm lies on the very fast elucidation of differentiation between sample classes and the identification of segments in the chromatogram which are responsible for differences between classes. However, to provide confirmation of the importance of the found segments, a more detailed downstream analysis for the wine data set is provided. Subsequently, with the rice and the urea data set two further examples of the performance of the new approach are provided focussing on the comparison of the results of class differentiation of our approach to the published data.

3.1. Wine dataset

The wine dataset consists of 39 chromatograms of wines fermented with three different commercial yeast products. Number of samples for each class were 20, 12 and 7 for Clos, Rbs and Vrb respectively [10]. No baseline correction or linear alignment was performed. TICs of all samples were summed to create a reference chromatogram for the segmentation, which resulted in 78 segments. Overall ranking of all segments is visualized by plotting the segment scores of the two metrics micro average of AUC and the maximum average value of the PR curve area against each other (Fig. S1).

After the ranking of the modeled segments according to the ability to differentiate between the three classes, 10 segments exceeded the cut-off values of the model evaluation metrics of 0.9 and were therefore selected for further investigation. A visual inspection of the TIC overlays of these segments provides clear evidence that these segments contain peaks, which clearly contribute to the differentiation of sample classes (Figs. S2–11). Examples of rejected and accepted segments are shown in Fig. 2. Segment loadings of the tucker decomposition of accepted segments are visualized using t-SNE. The iteration was stopped after no further progress was made with an error of 0.0474 after 119250 iterations. All three classes are clearly separated (Fig. 3). The class Rbs differentiates along the first dimension whereas the Vrb and Clos differentiate along the second dimension (Fig. 3). Moreover, an equivalent result figure is also found by using the approach described in the original work [10] which the reader is referred to.

Our proposed workflow is easily adjustable to further downstream processes such as PARADISE [64]. Moreover, in order to investigate deeper on the actual chemical compounds as potential biomarkers, PARAFAC2 deconvolution and annotation of deconvoluted mass spectra to NIST library of all peaks in the 10 accepted segments was performed as described in the reference method [10]. In total 15 peaks were found (Table S1). A comparison of the results with the original work reveals that 13 compounds were found by both studies while one unknown compound (LRI 1263) was not found with proposed supervised learning method. Note, that this unknown compound is contained in segment 31 which has still very high metrics scores of 0.89 and 0.78 for average

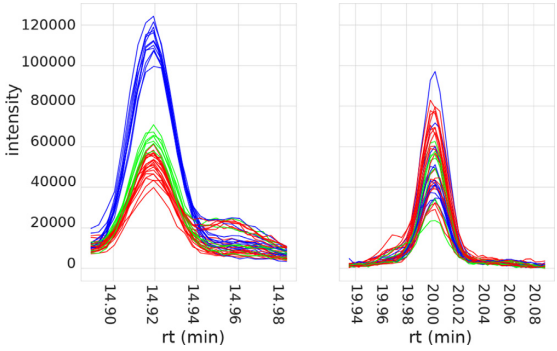


Fig. 2. TICs of a good and a poor segment found by the classification model of the wine dataset [10]. Good classification happens when one or multiple classes are clearly differentiated from the rest. For poor classification no separation between classes is observed. Colors correspond to the different commercial wine yeast starter cultures. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

precision score and micro average of ROC curve, respectively. Interestingly, two new compounds were found with the here proposed supervised learning method, which were not considered in the original paper (Table S1). These two compounds were tentatively identified as ethyl heptanoate and linalool. Particularly, linalool is known to be an important wine aroma constitute and the concentration of linalool can be strongly influenced by the yeast strain used during fermentation [65].

3.2. Urea dataset

The urease pretreatment dataset was downloaded from MetaboLights [54] with identifier MTBLS71 [55]. The experiment consists of 160 samples, with two separate classes: female-male; and no pretreatment - urease pretreated, yielding 4 classes in total. Although, this could be seen as a multi-labelled classification task, a multi-class approach was done in accordance with our approach.

TICs of all samples were summed to create a reference chromatogram which was used for segmentation resulting in 95 segments. 14 segments were retained for downstream analysis. Tucker decomposed loadings of the chosen segments were visualized using t-SNE with an error of 0.4928 after 300000 iterations. As the data was more noisy, the perplexity value was set to 4 from default 10 to emphasize more local effects [55]. The main separation was observed among urease pretreated and the control, while there seems to be a trend dividing the female-male class especially in the urease pretreated class (Fig. 4). The reference method gives a corresponding result figure as described in the original work [55] which the reader is referred to. Additionally, the visualization shows clearly that the experiment was done with biological duplicates. These findings are in accordance with the original work, where the main driving factor of variability was identified as the pretreatment while variability related to gender was only observed in lower-resolution [55].

3.3. Rice dataset

The rice cultivar dataset was downloaded from MetaboLights [54] with identifier MTBLS288 [56] and consists of 79 samples. Two sets of classes are defined: grain cultivars (4 classes) and development measured in days (5 classes). In regards to the original

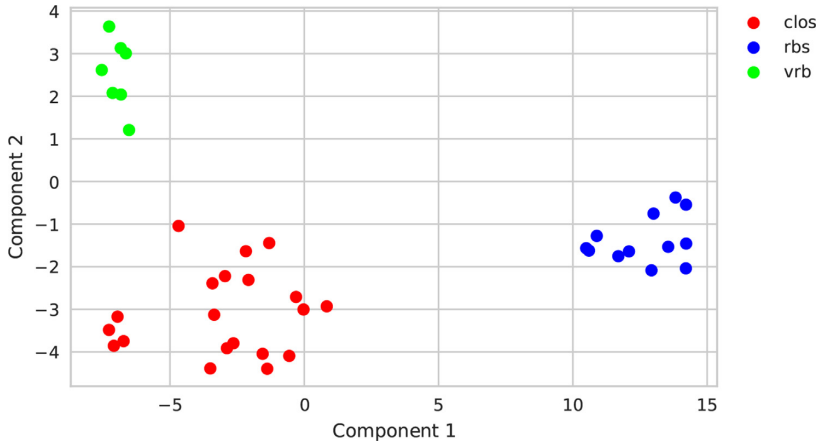


Fig. 3. t-SNE visualization of the wine dataset [10]. Tucker decomposed sample loadings of the ten highest scored segments are projected with two component t-SNE visualization. The three commercial wine yeast products are clearly separated from each other. Colors correspond to the different commercial wine yeast products. Rbs: commercial yeast product 1, Clos: commercial yeast product 2, Vrb: commercial yeast product 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

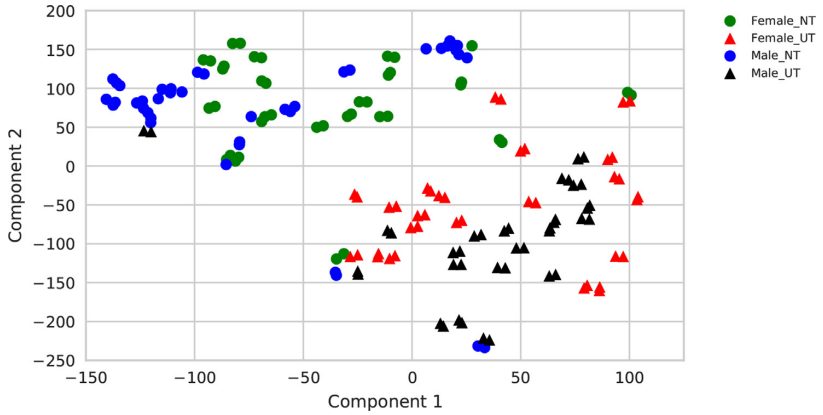


Fig. 4. t-SNE visualization of the urea dataset [55]. Tucker decomposed sample loadings of the 14 highest scored segments are projected with two component t-SNE visualization. Highest variation in the dataset is between the urease pretreated and no pretreated. Additionally, urease pretreated male and female classes seem to be different. NT: no pretreatment; UT: urease pretreatment.

authors work we chose to combine these two to create 20 classes with each having 4 samples. For the group 14_Nip only 3 samples were found due to some inconsistency in the public repository. No baseline correction or linear alignment was performed. TICs of all samples were summed to create a reference chromatogram which was used to for segmentation resulting in 134 segments. In total 38 segments were retained after running the pipeline. A comparably large number of important chemical compounds were also observed by the original authors [56].

Further, the sample loadings of the tucker decomposition of the retained segments were visualized with t-SNE. The iteration was stopped after no further progress was made with an error of 0.0697 after 105750 iterations. From the t-SNE representation it is evident that the rice grain development days form a gradient (Fig. 5). Early stage development samples (7–14 days) are seemingly more separated than the late stage samples (28–42 days). In the latter the cultivar effect seems to be more dominant (Fig. 5). These findings follow the same pattern as reported in the original work, in which the resolved metabolic features were presented through PCA [56] with a results figure which the reader is referred to. By taking overfitting into consideration, this example verifies that even for small sample sizes per class, it is still possible with our fully automated approach to extract the coincident and meaningful information as reported in the original work.

3.4. Impact of segmentation on algorithm performance

In order to investigate the effect of the robust segmentation on the results of the approach, four different segmentation scenarios were tested. Differing numbers of segments were achieved by varying the amount of RMS-noise for the detection of minima in the reference chromatogram. For all datasets it was found that the number of chosen segments is reaching a plateau once a relevant number of segments is reached (Tables S2–S5). For the urea dataset the segments with metric values larger than 0.9 after running our pipeline were deconvoluted with PARAFAC2 to obtain more

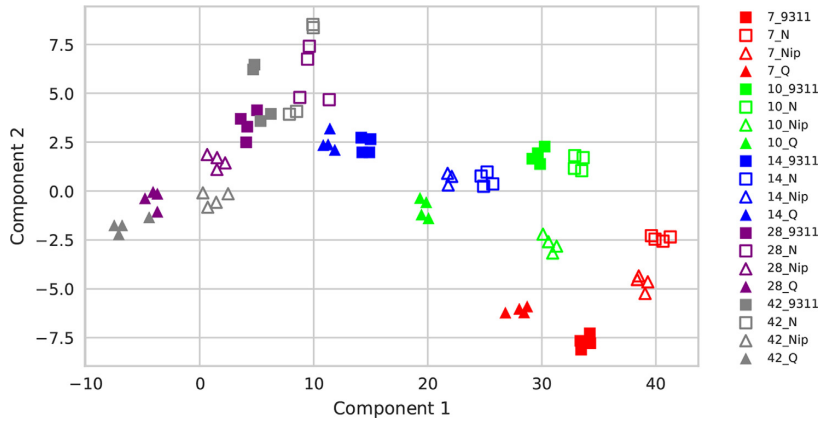


Fig. 5. t-SNE visualization of the rice dataset [56]. Tucker decomposed sample loadings of the 30 highest scored segments are projected with two component t-SNE visualization. The first component of the t-SNE distinguishes rice grains from different developmental stages. The early stage development (7-14) variation is higher than the late stage development (28−42), where the cultivar effect seems to be more dominant. The numbers in the sample names correspond to grain development measured in days. The labels 9311, N, Nip and O correspond to the cultivars 9311, Nongken 58, Nipponbare and Qingfengai, respectively.

information on the performance during downstream analysis. The selected deconvoluted peaks also reach a maximum number (Table S5). Furthermore, clustering performance, determined with cophenetic correlation coefficient, follows the same behavior as described above. Robust linear alignment leads to a small increased number of important deconvolute compounds (Table S5).

Smaller segment intervals are generally preferable as they are easier to deconvolute and to investigate visually. Large intervals with low number of segments are robust but also prone to lose information. While the classifier works on raw data signal (taking interactions into account) the evaluation of the classifier performance becomes less pronounced for complex segments, which in turn leads to lower segment selection metrics. For exploratory data analysis purposes segment selection parameter thresholds can be set lower in order to increase the amount of false positives, which can be identified and removed during the downstream analysis process. Small intervals with many segments lead however not necessarily to more important information in the downstream analysis as demonstrated by the cophenetic correlation table for the urea dataset (Table S5).

The size of segments is also crucial for the tensor decomposition and downstream processing, as the rank of the segment tensors is related to speed and accuracy of the model performance [66]. Smaller segments are easier to deconvolute for instance with PARAFAC2 to obtain more reliable and valid models [67]. Other interesting strategies for segmentation of chromatograms could be further included in our approach. Our data analysis strategy could also be used with other established methods for feature detection instead of using transformed chromatogram segments [4,22,67,68], albeit the advantages of modeling decomposed raw data segments would of course be lost. Using our approach with other established strategies for feature selection might not only increase resolution of the approach but makes the model applicable to other chromatographic techniques and to high resolution mass spectrometry.

3.5. Algorithmic efficiency and user interaction

Besides the optional control of the segmentation with desired level of RMS-noise and the inspection of results no further user interaction is needed for our approach. The calculation time of the algorithm depends on the total scan numbers in the chromatogram and the number of samples. The computations were performed on a computer equipped with an quad core Intel Core i7-6700 CPU with 3.4 Ghz. Total computation time including all segments for the three data sets took 1 min, 12.7 min and 20.7 min for the wine, urea and rice data set, respectively. Downstream analysis including repetition of the algorithm with only important segments took additional 0.1 min, 3.3 min and 2.5 min, respectively. These facts make our approach a very fast and strong tool to find class differentiation in non-targeted GC-MS data and reducing the time spent on total data analysis tremendously. Through greater efficiency and a shorter analysis route, specific important regions in the chromatograms can be discovered for any further investigation.

4. Conclusions

In this paper we introduce a novel, automated workflow for non-targeted GC-MS exploratory data analysis using supervised machine learning exploiting the power of sample size and knowledge of class memberships. Remarkably, with only segmented chromatographic raw data signals, meaningful information can be extracted in an automated way allowing the user productively focus on the downstreaming of only important regions in the chromatograms, which are responsible for class separation.

Our approach is reducing manipulation of data during the data analysis progress such as peak picking, deconvolution and retention time alignment. We have shown that our approach is able to reproduce the results of three published datasets. The key benefit of this automatized workflow is to speed up data analysis and facilitate differentiation between sample groups straight from the chromatographic raw data signals allowing the user to focus on the inspection of the relevant regions. Classification is therefore directly achieved, whereas additional tools for downstream analysis are necessary for the identification of biomarkers. The proposed workflow is freely available and can be found at https://github.com/kkpsiren/vesi/.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgments

Work by KS was funded by the Horizon 2020 Programme of the European Commission within the Marie Skłodowska-Curie Innovative Training Network “MicroWine” (grant number 643063). Work by JV was funded by the funding initiative "Plant Breeding Research for the Bioeconomy" under the National Research Strategy BioEconomy 2030 of the Federal Ministry of Education and Research (BMBF) and the Federal Ministry of Food and Agriculture (BMEL) of the Federal Republic of Germany (grant number 031B0197).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.acax.2019.100005.

References

[1] L. Yi, N. Dong, Y. Yun, B. Deng, D. Ren, S. Liu, Y. Liang, Chemometric methods in data processing of mass spectrometry-based metabolomics: a review, Anal. Chim. Acta 914 (2016) 17–34.
[2] J. Boccard, J.-L. Veuthey, S. Rudaz, Knowledge discovery in metabolomics: an overview of MS data handling, J. Separ. Sci. 33 (2010) 290–304.
[3] R. Goodacre, S. Vaidyanathan, W.B. Dunn, G.G. Harrigan, D.B. Kell, Metabolomics by numbers: acquiring and understanding global metabolite data, Trends Biotechnol. 22 (2004) 245–252.
[4] R. Tautenhahn, C. Böttcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, BMC Bioinf. 9 (2008) 504.
[5] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, Nat. Methods 12 (2015) 523–526.
[6] R. Spicer, R.M. Salek, P. Moreno, D. Cañueto, C. Steinbeck, Navigating freely-available software tools for metabolomics analysis, Metabolomics 13 (2017) 106.
[7] M.-E.P. Papadimitropoulos, C.G. Vasilopoulou, C. Maga-Nteve, M.I. Klapa, Untargeted GC-MS metabolomics, in: G.A. Theodoridis, H.G. Gika, I.D. Wilson (Eds.), Metabolic Profiling: Methods and Protocols, Springer New York, New York, NY, 2018, pp. 133–147.
[8] M.L. Santoru, C. Piras, A. Murgia, V. Palmas, T. Camboni, S. Liggi, I. Ibba, M.A. Lai, S. Orrù, S. Blois, A.L. Loizedda, J.L. Griffin, P. Usai, P. Caboni, L. Atzori, A. Manzin, Cross sectional evaluation of the gut-microbiome metabolome axis in an Italian cohort of IBD patients, Sci. Rep. 7 (2017) 9523.
[9] J.M. Amigo, T. Skov, R. Bro, J. Coello, S. Maspoch, Solving GC-MS problems with PARAFAC2, Trends Anal. Chem. 27 (2008) 714–725.
[10] J. Vestner, G. de Revel, S. Krieger-Weber, D. Rauhut, M. du Toit, A. de Villiers, Toward automated chromatographic fingerprinting: a non-alignment approach to gas chromatography mass spectrometry data, Anal. Chim. Acta 911 (2016) 42–58.
[11] J.H. Christensen, J. Mortensen, A.B. Hansen, O. Andersen, Chromatographic preprocessing of GC–MS data for analysis of complex chemical mixtures, J. Chromatogr., A 1062 (2005) 113–123.
[12] M. Cocchi, C. Durante, M. Grandi, D. Manzini, A. Marchetti, Three-way principal component analysis of the volatile fraction by HS-SPME/GC of aceto balsamico tradizionale of modena, Talanta 74 (2008) 547–554.
[13] M. Daszykowski, R. Danielsson, B. Walczak, No-alignment-strategies for exploring a set of two-way data tables obtained from capillary electrophoresis—mass spectrometry, J. Chromatogr., A 1192 (2008) 157–165.
[14] N.A. Sinkov, J.J. Harynuk, Cluster resolution: a metric for automated, objective and optimized feature selection in chemometric modeling, Talanta 83 (2011) 1079–1087.
[15] D. Ballabio, T. Skov, R. Leardi, R. Bro, Classification of GC-MS measurements of wines by combining data dimension reduction and variable selection techniques, J. Chemom. 22 (2008) 457–463.
[16] J. Jaumot, N. Escaja, R. Gargallo, C. González, E. Pedroso, R. Tauler, Multivariate curve resolution: a powerful tool for the analysis of conformational transitions in nucleic acids, Nucleic Acids Res. 30 (2002) e92.
[17] M. Daszykowski, B. Walczak, Methods for the exploratory analysis of two-dimensional chromatographic signals, Talanta 83 (2011) 1088–1097.
[18] R. Bro, C.A. Andersson, H.A.L. Kiers, PARAFAC2—Part II. Modeling chromatographic data with retention time shifts, J. Chemom. 13 (1999) 295–309.
[19] L.G. Johnsen, J.M. Amigo, T. Skov, R. Bro, Automated resolution of overlapping peaks in chromatographic data: chromatographic data analysis, J. Chemom. 28 (2014) 71–82.
[20] X. Domingo-Almenara, A. Perera, J. Brezmes, Avoiding hard chromatographic segmentation: a moving window approach for the automated resolution of gas chromatography-mass spectrometry-based metabolomics signals by multivariate methods, J. Chromatogr., A 1474 (2016) 145–151.
[21] X. Domingo-Almenara, J. Brezmes, M. Vinaixa, S. Samino, N. Ramirez, M. Ramon-Krauel, C. Lerin, M. Díaz, L. Ibáñez, X. Correig, A. Perera-Lluna, O. Yanes, eRah: a computational tool integrating spectral deconvolution and alignment with quantification and identification of metabolites in GC/MS-based metabolomics, Anal. Chem. 88 (2016) 9821–9829.
[22] A. Smirnov, W. Jia, D.I. Walker, D.P. Jones, X. Du, ADAP-GC 3.2: graphical software tool for efficient spectral deconvolution of gas chromatography-high-resolution mass spectrometry metabolomics data, J. Proteome Res. 17 (2018) 470–478.
[23] L.A. Adutwum, R.J. Abel, J. Harynuk, Total ion spectra versus segmented total ion spectra as preprocessing tools for gas chromatography - mass spectrometry data, J. Forensic Sci. 63 (2018) 1059–1068.
[24] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC Bioinf. 11 (2010) 395.
[25] L. Han, Y.-M. Zhang, J.-J. Song, M.-J. Fan, Y.-J. Yu, P.-P. Liu, Q.-X. Zheng, Q.-S. Chen, C.-C. Bai, T. Sun, Y.-B. She, Automatic untargeted metabolic profiling analysis coupled with Chemometrics for improving metabolite identification quality to enhance geographical origin discrimination capability, J. Chromatogr., A 1541 (2018) 12–20.
[26] Y.-J. Yu, H.-Y. Fu, L. Zhang, X.-Y. Wang, P.-J. Sun, X.-B. Zhang, F.-W. Xie, A chemometric-assisted method based on gas chromatography-mass spectrometry for metabolic profiling analysis, J. Chromatogr., A 1399 (2015) 65–73.
[27] A. Trimigno, L. Münger, G. Picone, C. Freiburghaus, G. Pimentel, N. Vionnet, F. Pralong, F. Capozzi, R. Badertscher, G. Vergères, GC-MS based metabolomics and NMR spectroscopy investigation of food intake biomarkers for milk and cheese in serum of healthy humans, Metabolites 8 (2018) 26.
[28] H. Nikpour, M. Mousavi, H. Asadollahzadeh, Qualitative and quantitative analysis of Teucrium polium essential oil components by GC-MS coupled with MCR and PARAFAC methods, Phytochem. Anal. 29 (2018) 590–600.
[29] C. Chen, J. Husny, S. Rabe, Predicting fishiness off-flavour and identifying compounds of lipid oxidation in dairy powders by SPME-GC/MS and machine learning, Int. Dairy J. 77 (2018) 19–28.
[30] S. Taghadomi-Saberi, S. Mas Garcia, A. Allah Masoumi, M. Sadeghi, S. Marco, Classification of bitter orange essential oils according to fruit ripening stage by untargeted chemical profiling and machine learning, Sensors 18 (2018) 1922.
[31] A. Acharjee, Z. Ament, J.A. West, E. Stanley, J.L. Griffin, Integration of metabolomics, lipidomics and clinical data using a machine learning method, BMC Bioinf. 17 (2016) 440.
[32] Q. Yang, L. Xu, L.-J. Tang, J.-T. Yang, B.-Q. Wu, N. Chen, J.-H. Jiang, R.-Q. Yu, Simultaneous detection of multiple inherited metabolic diseases using GC-MS urinary metabolomics by chemometrics multi-class classification strategies, Talanta 186 (2018) 489–496.
[33] A. Smolinska, A.-C. Hauschild, R.R.R. Fijten, J.W. Dallinga, J. Baumbach, F.J. van Schooten, Current breathomics–a review on data pre-processing techniques and machine learning in metabolomics breath analysis, J. Breath Res. 8 (2014) 027105.
[34] X. Wang, D. Wang, J. Wu, X. Yu, J. Lv, J. Kong, G. Zhu, R. Su, Metabolic characterization of myocardial infarction using GC-MS-based tissue metabolomics, Int. Heart J. 58 (2017) 441–446.
[35] Z. Lai, H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, A. Ogiwara, J. Meissen, M. Showalter, K. Takeuchi, T. Kind, P. Beal, M. Arita, O. Fiehn, Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics, Nat. Methods 15 (2018) 53–56.
[36] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (1966) 279–311.
[37] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY, USA, 2016, pp. 785–794.
[38] T. Fawcett, An introduction to ROC analysis, Pattern Recogn. Lett. 27 (2006) 861–874.
[39] M. Sokolova, N. Japkowicz, S. Szpakowicz, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: Sattar A., Kang B. (eds) AI 2006: Advances in Artificial Intelligence. AI 2006. Lecture Notes in Computer Science, vol. 4304. Springer, Berlin, Heidelberg.
[40] J. Ambroise, A. Robert, B. Macq, J.-L. Gala, Transcriptional network inference from functional similarity and expression data: a global supervised approach, Stat. Appl. Genet. Mol. Biol. 11 (2012). Article 2.
[41] J. Davis, M. Goadrich, The relationship between Precision-Recall and ROC curves, in: Proceedings of the 23rd International Conference, 2006. https://www.biostat.wisc.edu/~page/rocpr.pdf. (Accessed 6 July 2018).
[42] P.D. Foster Provost, Well-Trained PETs: Improving Probability Estimation Trees, 2000. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.33.309. (Accessed 6 July 2018).
[43] M. Carbonero-Ruz, F.J. Martínez-Estudillo, F. Fernández-Navarro, D. Becerra-Alonso, A.C. Martínez-Estudillo, A two dimensional accuracy-based measure for classification performance, Inf. Sci. 382–383 (2017) 60–80.
[44] L. van der Maaten, G. Hinton, Visualizing Data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.

[45] K. Tian, L. Wu, S. Min, R. Bro, Geometric search: a new approach for fitting PARAFAC2 models on GC-MS data, Talanta 185 (2018) 378–386.
[46] P. Software Foundation, Python Language Reference, 2017, Version 3.6. 3.
[47] E. Jones, T. Oliphant, P. Peterson, {SciPy}: Open Source Scientific Tools for {Python}, 2001. http://www.scipy.org.
[48] T. Kolda, B. Bader, Tensor decompositions and applications, SIAM Rev. 51 (2009) 455–500.
[49] J. Kossaifi, Y. Panagakis, A. Anandkumar, M. Pantic, TensorLy: Tensor Learning in Python, CoRR, 2018. http://arxiv.org/abs/1610.09555.
[50] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, Others, scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[51] L. van der Maaten, Accelerating t-SNE using tree-based algorithms, J. Mach. Learn. Res. 15 (2014) 3221–3245.
[52] J.D. Hunter, Matplotlib: a 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95.
[53] M. Waskom, O. Botvinnik, D. O'Kane, P. Hobson, S. Lukauskas, D.C. Gemperline, T. Augspurger, Y. Halchenko, J.B. Cole, J. Warmenhoven, J. de Ruiter, C. Pye, S. Hoyer, J. Vanderplas, S. Villalba, G. Kunter, E. Quintero, P. Bachant, M. Martin, K. Meyer, A. Miles, Y. Ram, T. Yarkoni, M.L. Williams, C. Evans, C. Fitzgerald, Brian, C. Fonnesbeck, A. Lee, A. Qalieh, mwaskom/seaborn: v0.8.1 (September 2017), 2017, https://doi.org/10.5281/zenodo.883859.
[54] K. Haug, R.M. Salek, P. Conesa, J. Hastings, P. de Matos, M. Rijnbeek, T. Mahendraker, M. Williams, S. Neumann, P. Rocca-Serra, E. Maguire, A. González-Beltrán, S.-A. Sansone, J.L. Griffin, C. Steinbeck, MetaboLights–an open-access general-purpose repository for metabolomics studies and associated meta-data, Nucleic Acids Res. 41 (2013) D781–D786.
[55] B.-J. Webb-Robertson, Y.-M. Kim, E.M. Zink, K.A. Hallaian, Q. Zhang, R. Madupu, K.M. Waters, T.O. Metz, A statistical analysis of the effects of urease pre-treatment on the measurement of the urinary metabolome by gas chromatography-mass spectrometry, Metabolomics 10 (2014) 897–908.
[56] C. Hu, T. Tohge, S.-A. Chan, Y. Song, J. Rao, B. Cui, H. Lin, L. Wang, A.R. Fernie, D. Zhang, J. Shi, Identification of conserved and diverse metabolic shifts during rice grain development, Sci. Rep. 6 (2016) 20942.
[57] X. Liu, X.-H. Zhu, P. Qiu, W. Chen, A correlation-matrix-based hierarchical clustering method for functional connectivity analysis, J. Neurosci. Methods 211 (2012) 94–102.
[58] D. Müllner, fastcluster: fast hierarchical, agglomerative clustering routines for R and Python, J. Stat. Software (2013). https://www.jstatsoft.org/article/view/v053i09/v53i09.pdf.
[59] G. Bradski, The opencv library, Dr. Dobb's Journal of Software Tools (2000), 2000.
[60] N.E. Helwig, Multiway: Component Models for Multi-Way Data, 2018. https://CRAN.R-project.org/package=multiway.
[61] R Core Team, A Language and Environment for Statistical Computing, 2017. https://www.R-project.org/.
[62] A. Liaw, M. Wiener, Classification and regression by randomForest, R. News 2 (2002) 18–22. http://CRAN.R-project.org/doc/Rnews/.
[63] M.B. Kursa, W.R. Rudnicki, Others, feature selection with the boruta package, J. Stat. Software 36 (2010) 1–13.
[64] L.G. Johnsen, P.B. Skou, B. Khakimov, R. Bro, Gas chromatography–mass spectrometry data processing made easy, J. Chromatogr., A 1503 (2017) 57–64.
[65] A. Rapp, H. Mandery, Wine aroma, Experientia 42 (1986) 873–884.
[66] I. Oseledets, D. Savostianov, E. Tyrtyshnikov, Tucker dimensionality reduction of three-dimensional arrays in linear time, SIAM J. Matrix Anal. Appl. 30 (2008) 939–956.
[67] B. Khakimov, R.J. Mongi, K.M. Sørensen, B.K. Ndabikunze, B.E. Chove, S.B. Engelsen, A comprehensive and comparative GC-MS metabolomics study of non-volatiles in Tanzanian grown mango, pineapple, jackfruit, baobab and tamarind fruits, Food Chem. 213 (2016) 691–699.
[68] N. Dalmau, C. Bedia, R. Tauler, Validation of the regions of interest multivariate curve resolution (ROIMCR) procedure for untargeted LC-MS lipidomic analysis, Anal. Chim. Acta 1025 (2018) 80–91.

Supplementary information of Chapter 2

Supporting Information for

Automated supervised learning pipeline for non-targeted GC-MS data analysis

Kimmo Sirén [a,b], Ulrich Fischer [a], Jochen Vestner [a,*]
[a] Institute for Viticulture and Oenology, DLR Rheinpfalz, Breitenweg 71, D-67435 Neustadt, Germany
[b] Department of Chemistry, University of Kaiserslautern, Erwin-Schroedinger-Strasse 52, D-67663 Kaiserslautern Germany

*corresponding author.
Email: jochen.vestner@dlr-rlp.de
Phone: +49 6321671459

Table of Contents

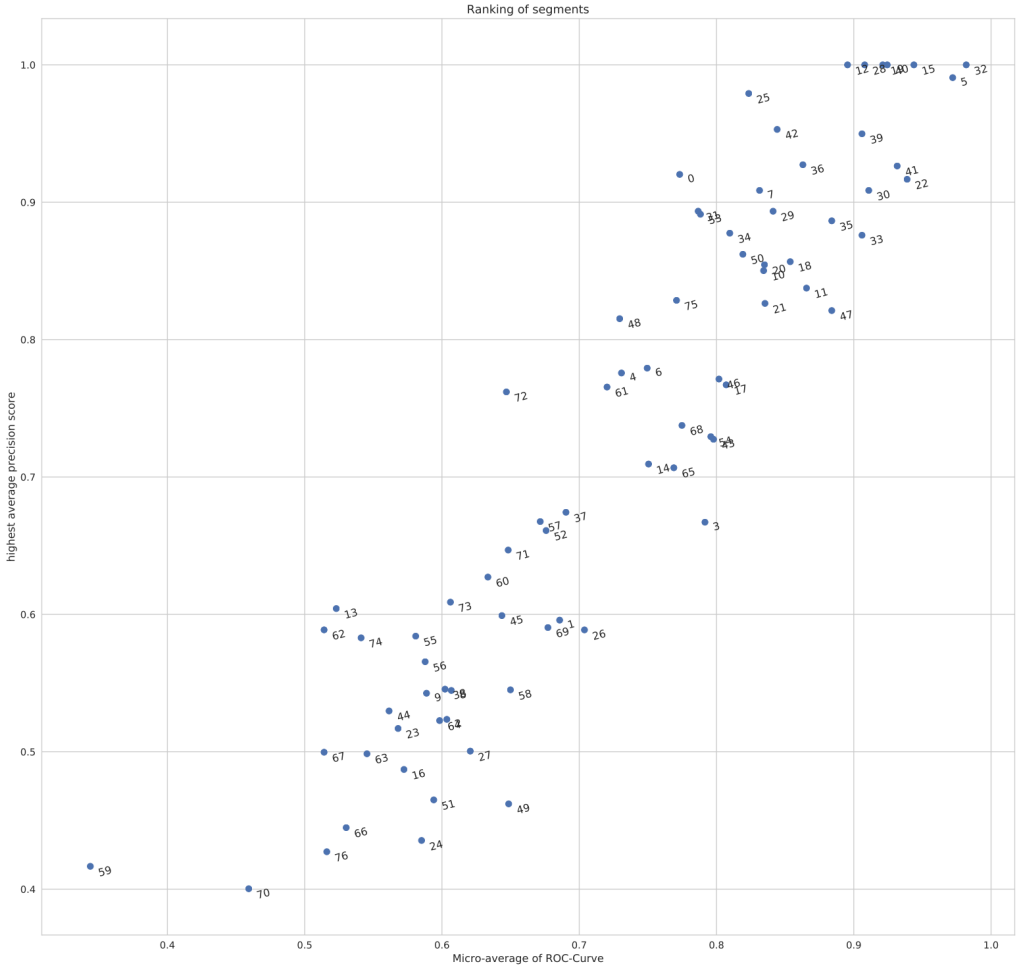
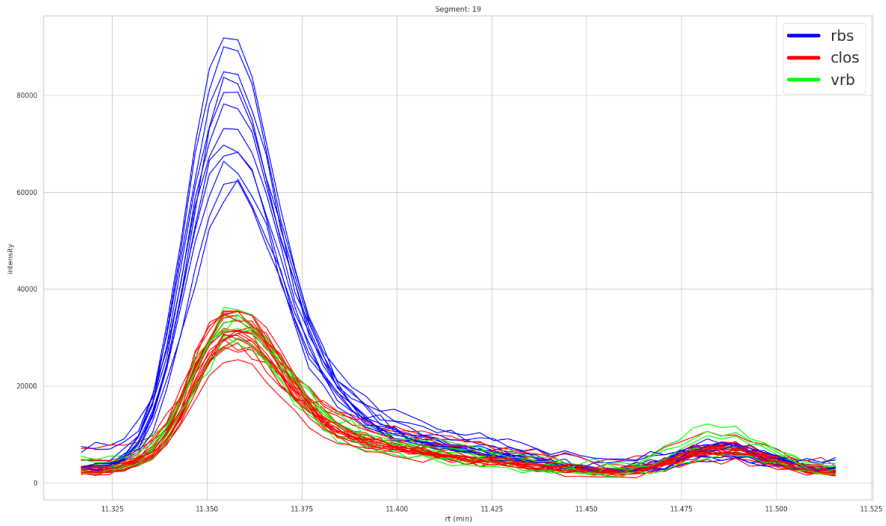
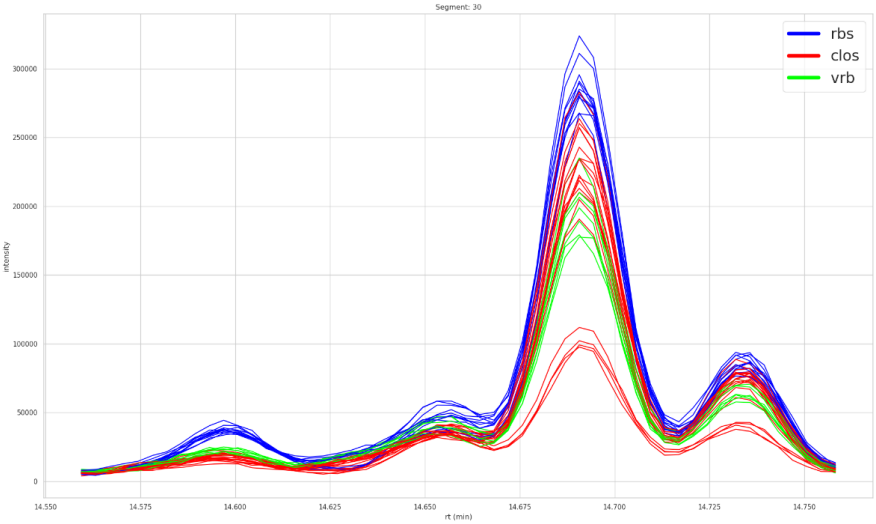
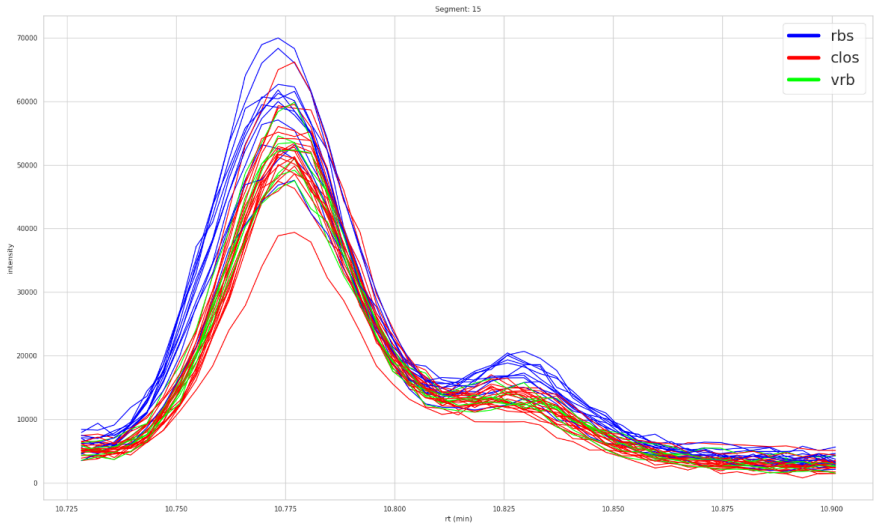
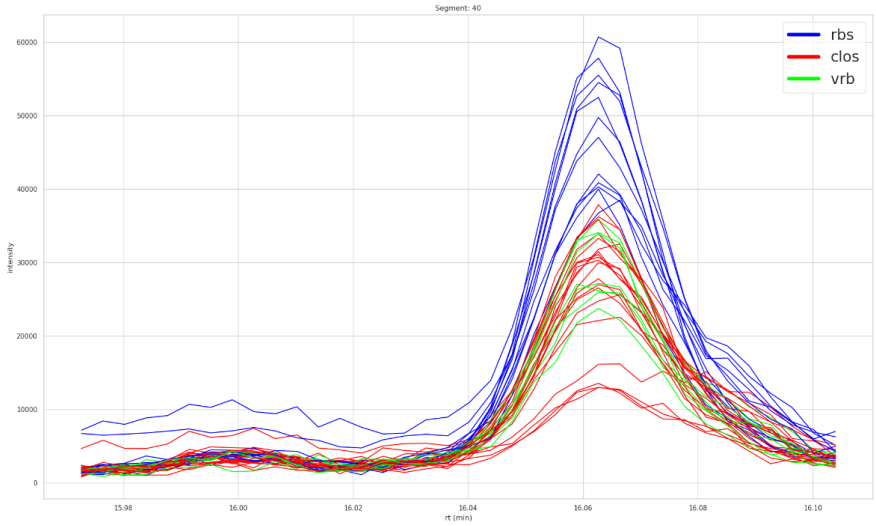


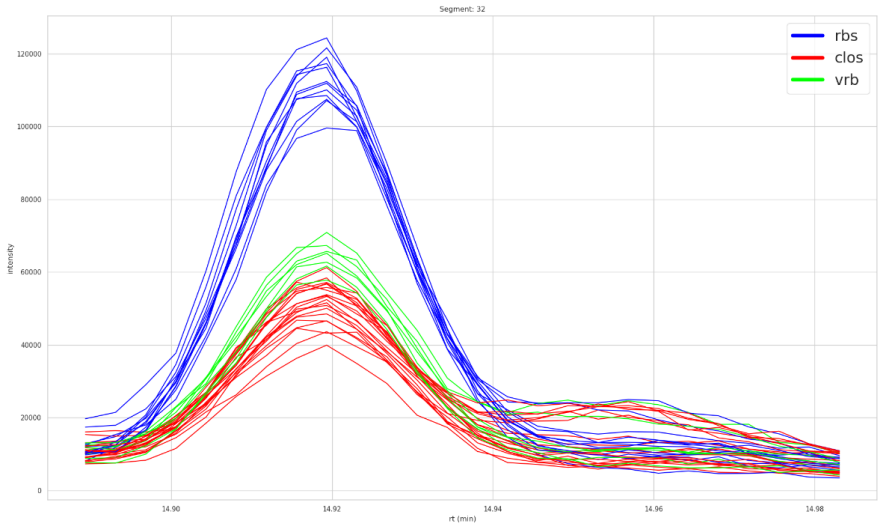
Figure S1. Ranking of segments in the wine dataset. Segments are sorted based on the micro-average of ROC Curve threshold and maximum average values of all classes for Precision Recall curve class. Retained segments for downstream analysis were required to have both values higher than 0.90.

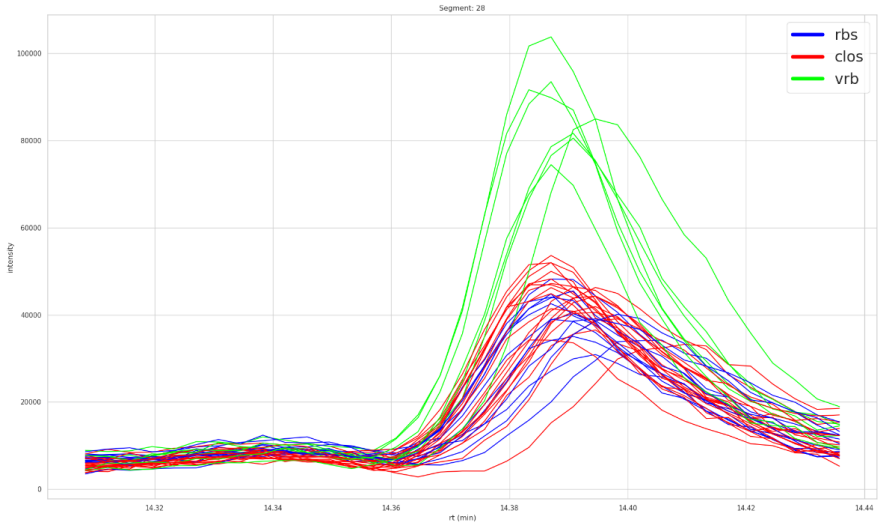

Segment: 19
rbs
clos
vrb
intensity
rt (min)

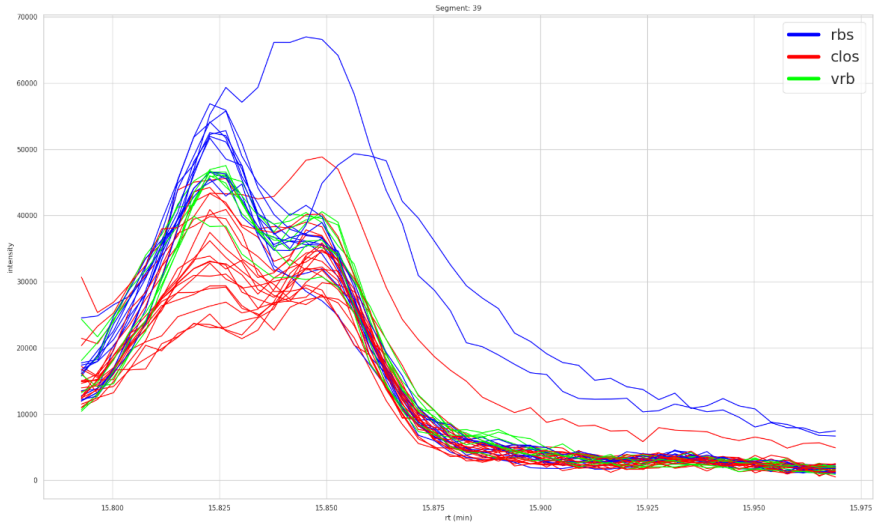

Segment: 30
rbs
clos
vrb
intensity
rt (min)

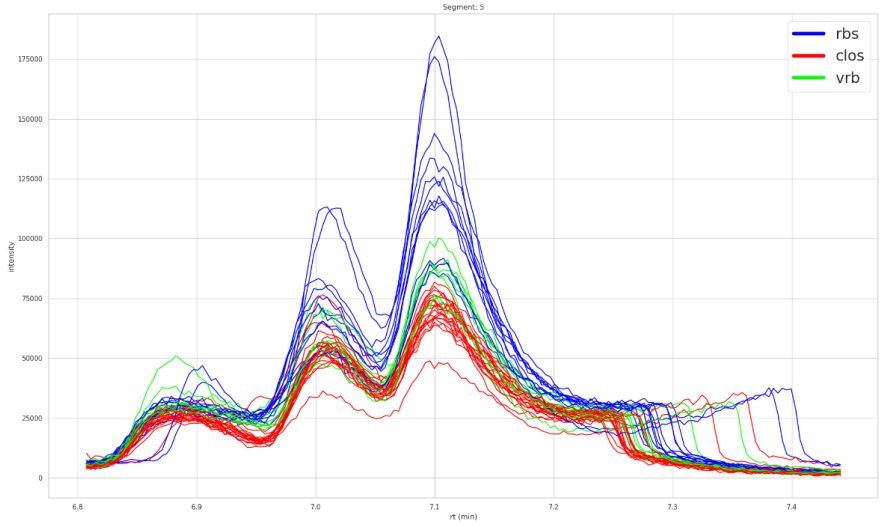

Segment: 15
rbs
clos
vrb
intensity
rt (min)
0
10000
20000
30000
40000
50000
60000
70000
10.725
10.750
10.775
10.800
10.825
10.850
10.875
10.900

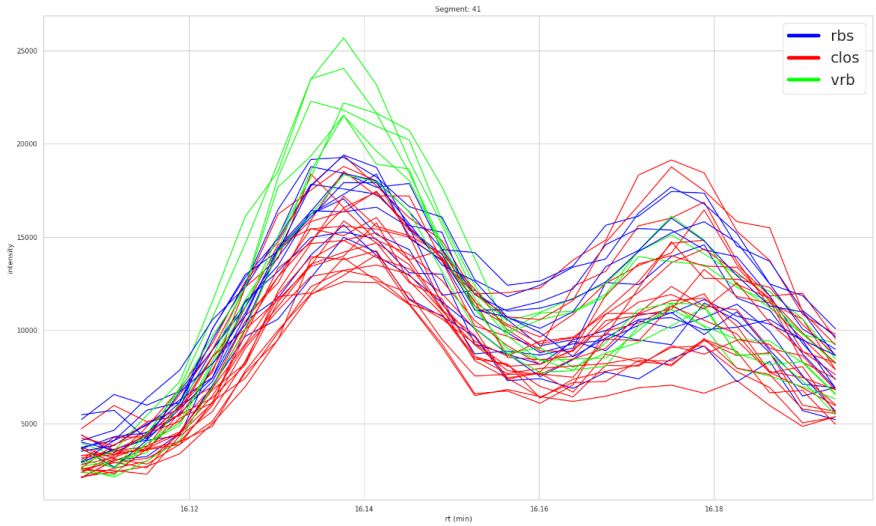

Segment: 40
rbs
clos
vrb
intensity
rt (min)
0
10000
20000
30000
40000
50000
60000
15.98
16.00
16.02
16.04
16.06
16.08
16.10

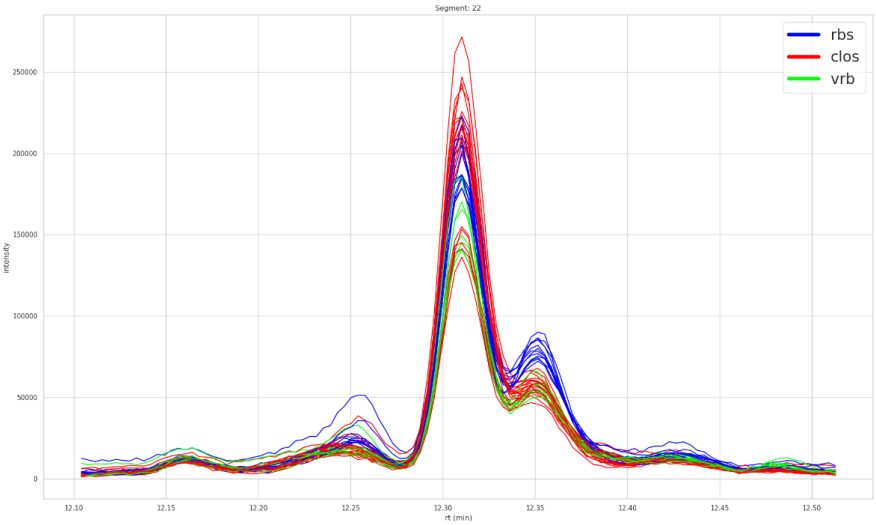

Segment: 32
rbs
clos
vrb
intensity
rt (min)


Segment: 28
rbs
clos
vrb
intensity
rt (min)


Segment: 39
rbs
clos
vrb
intensity
rt (min)


Segment: 5
rbs
clos
vrb
intensity
rt (min)





Figures S2-11. TICs of segments in the wine dataset with values of both metrics > 0.9. Segments are sorted according to their importance as chosen by the supervised learning algorithm.

Table S1.. List of potential biomarkers that were found significantly varying between three yeast starter cultures in the wine dataset. Only accepted segments with values for both metrics > 0.9 were investigated. Supervised approach list the segment id-number, x corresponds to found significant in the reference method [10].

Compound name	LRI	Supervised approach	Reference method [10]
butanoic acid, 2-methyl ester	857	Segment 5	x
Propanoic acid 3-methyl-, ethyl ester	1003	Segment 15	x
2-hexenoic acid, ethyl ester	1048	Segment 19	x
heptanoic acid ethyl ester	1098	Segment 22	not found
1,6-Octadien-3-ol, 3,7-dimethyl- (linalool)	1101	Segment 22	not found
unknown compound	1111	Segment 22	x
6-octen-1-ol, 3,7-dimethyl- (citronellol)	1231	Segment 28	x
unknown (terpenoid-like MS)	1246	Segment 30	x
benzeneacetic acid, ethyl ester (ethyl benzeneacetate)	1250	Segment 30	x
hexanoic acid, 3-methylbutyl ester (isopentyl hexanoate)	1252	Segment 30	x
hexanoic acid, 2-methylbutyl ester (2-methylbutyl	1255	Segment 30	x
unknown	1263	not found	x
unknown (terpenoid-like MS)	1267	Segment 32	x
unknown (succinic acid ester)	1331	Segment 39	x
octanoic acid, 2-methylpropyl ester (isobutyl octanoate)	1350	Segment 40	x
unknown	1352	Segment 41	x

Table S2. Impact of segmentation on algorithm performance for the wine dataset without linear alignment.

RMS-noise	Total segments	Chosen segments	Cophenetic correlation
100	28	19	0.865
10	78	33	0.889
1	134	40	0.893
0.1	189	42	0.908

Table S3. Impact of segmentation on algorithm performance for the rice dataset without linear alignment.

RMS-noise	Total segments	Chosen segments	Cophenetic correlation
100	15	6	0.808
10	43	9	0.877
1	78	11	0.961
0.1	107	13	0.939

Table S4. Impact of segmentation on algorithm performance for the urea dataset without linear alignment and deconvoluted compounds in the chosen segments.

RMS-noise	Total segments	Chosen segments	Cophenetic correlation chosen segments	Deconvoluted compounds	Cophenetic correlation deconv. compounds
100	45	19	0.851	32	0.934
10	76	38	0.809	47	0.935
1	95	45	0.823	56	0.933
0.1	109	46	0.818	55	0.935

Table S5. Segments performance for the urea dataset with robust linear alignment and deconvoluted compounds in the chosen segments.

RMS-noise	Total segments	Chosen segments	Cophenetic correlation chosen segments	Deconvoluted compounds	Cophenetic correlation deconv. compounds
100	54	28	0.791	41	0.938
10	88	48	0.796	56	0.935
1	104	50	0.815	63	0.904
0.1	118	51	0.816	63	0.925

Chapter 3

Comparison of DNA extraction methods to detect fungal diversity in Riesling during alcoholic fermentation

Comparison of DNA extraction methods for use in fungal diversity analyses of Riesling during alcoholic fermentation

Sarah Siu Tze Mak[1§], Kimmo Sirén[2, 3§], Christian Carøe[1], Shyam Gopalakrishnan[1], Martin René Ellegaard[1], Franziska Klincke[4], Anders J. Hansen[5] , Ulrich Fischer[2], M. Thomas P. Gilbert[1*]

[1] Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

[2] Institute for Viticulture and Oenology, Dienstleistungszentrum Ländlicher Raum Rheinpfalz, Neustadt an der Weinstraße, Germany

[3] Department of Chemistry, University of Kaiserslautern, Kaiserslautern, Germany

[4] Department of Bio and Health Informatics, Technical University of Denmark, Kongens, Lyngby, Denmark

[5] Section for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

[§]authors contributed equally

* corresponding author; email: tgilbert@snm.ku.dk

Abstract

Recent advances in high-throughput sequencing (HTS) technologies allow us to obtain vast amounts of microbial information from wine and must samples. It is becoming accepted that approaches aimed at characterising the microbial diversity during fermentation can be enhanced, or possibly even replaced, with HTS-based metabarcoding. In this context, a number of different DNA extraction methods have been tested. To reduce experimental biases and increase data reproducibility, we compared three DNA extraction methods by evaluating differences in the fungal diversity, using 32 Riesling alcoholic fermentation samples originating from four vineyards (a total of 96 extracts). These extraction methods consisted of a commercial kit (method F), a recently published protocol that includes a DNA enhancer (method G), and a protocol based on a buffer containing common inhibitor removal reagents (method C). Fungal diversity was profiled using metabarcoding of two genetic markers, ITS2 and D2. Although all methods were able to distinguish vineyard effects on the fungal microbiomes, they differed quantitatively. In particular, while method G gave the highest DNA yields, method F yielded the highest diversity of OTUs. Overall, the results of this study highlight the importance of standardizing DNA extraction methods when characterising fungal diversity from wine and related samples.

Highlights

- We compared three different methods for extracting DNA from fermenting grape must
- We found quantitative differences in DNA yield and microbial sequences recovered
- However all three methods recovered vineyard specific variation in fungal diversity

Keywords

Wine; metabarcoding; high-throughput sequencing; vineyard; yeast diversity

Abbreviations

CTAB: Hexadecyltrimethylammonium bromide

EDTA: Ethylenediaminetetraacetic acid

F-value: pseudo-F value

HTS: High-throughput sequencing

PVP: polyvinylpyrrolidone

PVPP: polyvinylpolypyrrolidone

PCR: Polymerase chain reaction

TET: Tris-HCl, EDTA and Tween 20

R^2: coefficient of determination

1. Introduction

The dynamic process of transforming grape must into wine is mainly driven by activity of yeasts, in particular *Saccharomyces cerevisiae,* that is required for successful alcoholic fermentation. However, a wide diversity of other microbes enters the process, either because they are naturally found on grape berries (Bokulich et al., 2014), or within the viticultural and winery environments (Bokulich et al., 2013; Morrison-Whittle and Goddard, 2018). These include a number of non-*Saccharomyces* yeasts that are known to contribute to wine aroma and flavor (Bisson et al., 2017), such as members of the genera *Hanseniaspora*, *Lachancea*, *Metschnikowia*, *Pichia*, *Starmerella*, *Torulaspora* and *Wickerhamomyces*. The effect of these yeasts is not always desirable, with some of them known for their negative impact on wine quality (Belda et al., 2016a; Masneuf-Pomarede et al., 2015). Therefore, in order to control, predict and manipulate the outcome of the complex microbial processes underlying winemaking, it is important to understand which microbes are present and how they interact. This has been studied previously with isolation by culturing, followed by enzymatic screening and characterization (Belda et al., 2016b). Today, thanks to the development of increasingly economical high-throughput sequencing (HTS) techniques, it is becoming accepted that extant approaches aimed at characterising the microbial diversity during fermentation can be complemented, or possibly even replaced, with HTS-based metabarcoding (Lleixà et al., 2018; Sirén et al., 2019). Thus, there is a critical need to develop, standardise and perform comparisons of relevant methodologies and workflows, in order to maximise the reproducibility of such studies.

At the core of any HTS-based investigation is access to high yields of purified nucleic

acids. Furthermore, it has been repeatedly shown that the application of different extraction methods on the same microbiome datasets can produce different results (Dopheide et al., 2018; Inceoglu et al., 2010), in particular when substrates are complex, as is typically the case for fermented food samples (Keisam et al., 2016). This might relate to the efficacy of steps to remove inhibitors from the substrate - a problem often dealt with by incorporating extra purification steps in the extraction procedure (Lever et al., 2015). This is of particular relevance to those interested in the wine microbiome, as samples related to the fermentation process, from grape must to the final wine, can contain large quantities of tannins and other phenolic compounds, as well as polysaccharides (Jara et al., 2008; Siret et al., 2000), all of which may inhibit downstream enzymatic manipulation of the DNA (Cordero-Bueso et al., 2017; Jara et al., 2008; Lever et al., 2015). Given this, several previous studies have developed and compared methods aimed at recovering high quality purified DNA from must/wine samples, although principally with a focus on studying the DNA derived from the grape itself (review by Villano et al., 2017). In contrast, only two studies have explored the performance of different extraction methods on the microbes themselves, specifically one on acetic acid bacteria (Jara et al., 2008) and another on *S. cerevisiae* (Savazzini and Martinelli, 2006). Neither of these prior comparisons explored the ability to extract and isolate DNA from the wider microbial community, and both were applied to cultured samples. Thus, currently no optimised extraction protocol for wine exists. Nevertheless, a number of different extraction methods have been applied to must and wine samples. These vary from “handcrafted” protocols that include various inhibitor removal reagents (e.g. lysozyme, CTAB, PVP & PVPP, phenol-chloroform, etc.) (Işçi et al., 2014; Rodríguez-Plaza et al., 2006; Zepeda-Mendoza et al., 2018) to commercial extraction kits (Bokulich

et al., 2014; Godálová et al., 2016; Piao et al., 2015; Sternes et al., 2017). Therefore, comparisons of relevant DNA extraction methods are needed in order to improve the reproducibility of HTS-based wine studies.

Here, we compare the performance of three DNA extraction protocols of relevance to metabarcoding investigations of wine fungal diversity, using vinification samples derived from Riesling must obtained at four vineyards. The first method used a lysis buffer containing the common inhibitor removal reagents (modified from (Rodríguez-Plaza et al., 2006), while the second included a recently developed DNA enhancer called 'G2' (Zepeda-Mendoza et al., 2018). The third protocol was a commercial kit used in several previous studies, notably with Riesling (Piao et al., 2015; Stefanini et al., 2016). To compare the methods, we generated metabarcoding data of two fungal markers, one, the internal transcribed spacer 2 (ITS2) and the second, the D1/D2 domain of 26S of the rRNA gene. We evaluated the extraction protocols based on DNA concentration recovered, qPCR measured ITS2 and D2 gene copy numbers, fungal diversity, and each method's ease of use, financial cost and handling time.

2. Materials and methods

2.1 Sample collection and fermentation trial set up

Riesling grapes from four vineyards in Pfalz wine region in Germany were harvested directly into autoclaved sterile flat plastic bags (Neolab) in Autumn 2015. Bags were sealed in the vineyard and transported back to the Institute for Viticulture and Oenologie, DLR Rheinpfalz, Neustadt an der Weinstraße, Germany. Further processing was done under a laminar flow hood in sterile conditions. Grapes were crushed and pressed into 3L

sterile autoclaved Erlenmeyer flasks (Duran, Germany) for overnight sedimentation at 4 °C. The supernatant was transferred into duplicate 1 L autoclaved Erlenmeyer flasks (Duran, Germany) and sealed with silicone bungs attached to airlocks filled with distilled water. Samples were collected over four different days, where 4.5 mL of fermenting must was transferred into 5 mL DNA LoBind tubes (Eppendorf). Details of samples are shown in Table 1.

Table 1. Details of ferment samples analysed. Each sample name can be used to identify the vineyard name (first 2 characters), biological replicate (a or b), extraction method (abbreviation X replaced as either 1: method F; 2: method C & 3: method G), and the sampling date in 2015 (last 2 numbers).

No.	Sample name	Vineyard	Biological replicate	DNA concentration (ng/ μL) in different extraction methods			Sampling date
				F*	C	G	
1	w2aX09	2	a	< 0.02	1.34	0.51	09-Oct
2	w2bX09	2	b	0.28	0.70	0.76	09-Oct
3	w3aX09	3	a	0.21	8.72	1.30	09-Oct
4	w3bX09	3	b	0.45	2.97	5.30	09-Oct
5	w4aX09	4	a	0.12	0.41	0.30	09-Oct
6	w4bX09	4	b	0.12	0.26	0.32	09-Oct
7	w5aX09	5	a	< 0.02	0.59	0.18	09-Oct
8	w5bX09	5	b	< 0.02	0.21	0.31	09-Oct
9	w2aX12	2	a	0.22	1.19	0.21	12-Oct
10	w2bX12	2	b	0.29	0.87	1.14	12-Oct
11	w3aX12	3	a	0.41	2.80	3.54	12-Oct
12	w3bX12	3	b	0.29	0.96	5.96	12-Oct
13	w4aX12	4	a	0.52	5.26	4.45	12-Oct

14	w4bX12	4	b	0.14	0.41	0.60	12-Oct
15	w5aX12	5	a	< 0.02	0.32	0.27	12-Oct
16	w5bX12	5	b	0.36	0.86	1.94	12-Oct
17	w2aX14	2	a	< 0.02	0.88	0.23	14-Oct
18	w2bX14	2	b	0.30	0.93	1.43	14-Oct
19	w3aX14	3	a	0.19	2.94	2.69	14-Oct
20	w3bX14	3	b	0.30	0.69	1.58	14-Oct
21	w4aX14	4	a	0.47	3.92	4.04	14-Oct
22	w4bX14	4	b	0.18	0.48	0.51	14-Oct
23	w5aX14	5	a	0.21	3.61	1.19	14-Oct
24	w5bX14	5	b	0.5	2.03	3.70	14-Oct
25	w2aX20	2	a	0.33	3.48	3.48	20-Oct
26	w2bX20	2	b	0.23	0.65	0.85	20-Oct
27	w3aX20	3	a	< 0.02	0.28	0.13	20-Oct
28	w3bX20	3	b	< 0.02	0.23	0.36	20-Oct
29	w4aX20	4	a	0.22	0.30	0.21	20-Oct
30	w4bX20	4	b	0.11	0.18	0.53	20-Oct
31	w5aX20	5	a	< 0.02	0.76	0.75	20-Oct
32	w5bX20	5	b	0.28	0.20	1.06	20-Oct

*Concentration measured by Qubit HS dsDNA assay after cleaning up using DNA Clean and Concentrator™-5. Value "< 0.02 ng/μL" referred to concentrations below detection limit.

2.2 DNA extractions and qPCR

The three DNA extraction protocols varied by cell lysis and/or DNA purification procedures, ease of use, handling time and cost (Table 2). In total 32 samples (4.5 mL each) were extracted using each protocol. These represent biological replicates from 4 wineries sampled over 4 sampling dates, giving a total of 96 extracts (Table 1). All

extractions were performed once. Prior to DNA extraction, all samples were centrifuged at 6300 rpm for 10 min, after which the supernatant was removed and the pelleted cells resuspended in 1 mL of ice-cold 1× PBS (pH 7.4, Life Technologies, California, U.S.A.). This centrifugation and resuspension was repeated twice to maximise the debris removed. The final resuspended pellets were stored at -20 °C until extractions. All DNA extracts were quantified using a Qubit 1.0 fluorometer with dsDNA High Sensitivity Assay kit (ThermoFisher Scientific). Extraction blanks were included for every 14-16 samples, for each extraction method. More details on the methods follow.

2.2.1. Method F – Commercial kits

DNA was extracted using the FastDNA™ Spin Kit for Soil (MP Biomedical, California, U.S.A.) following the manufacturer's protocol, albeit with some modifications. Specifically, pellets were bead-beaten twice at 30 Hz for 40 seconds using a TissueLyser II (Qiagen, Hilden, Germany), with cooling on ice for 2 minutes in between bead-beating steps. In the final elution, 105 µL of 1× TET buffer, i.e. 1× TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0, Sigma-Aldrich) and 0.05% Tween 20 (Sigma-Aldrich) was added and incubated at 55 °C for 5 minutes before elution. Post extraction, the DNA was further concentrated from each sample using a DNA Clean and Concentrator™-5 (Zymo Research, California, U.S.), prior to final elution in 55 µL of 1× TET buffer. The whole process took ca. 1 hour 15 minutes per 10 samples.

2.2.2. Method C – CTAB/PVPP-chloroform protocol

Each pellet was added to a 2 mL DNA LoBind tube (Eppendorf) containing 50 µL 0.1

mm zirconia/silica beads (BioSpec Products, U.S.A.), then lysed in 300 µL of extraction buffer (modified from (Rodríguez-Plaza et al., 2006) containing: 2% Hexadecyltrimethylammonium bromide (CTAB, Sigma-Aldrich), 0.5% polyvinylpolypyrrolidone (PVPP, Sigma-Aldrich), 1% 2-mercaptoethanol, 20 mM EDTA (UltraPure 0.5 M EDTA, pH 8.0, Invitrogen), 10 mM Tris-HCl (1M, pH 8.0), 1.4 M sodium chloride (5M NaCl, ThermoFisher Scientific), Tween 20 (Sigma-Aldrich) and filled up with AccuGene molecular biology water (Lonza). Tubes were bead-beaten at 30 Hz for 2 minutes using a TissueLyser II (Qiagen, Hilden, Germany). Then, 3 µL Proteinase K (Qiagen, California, U.S.A.) was added to each tube and incubated at 45 °C for 60 minutes with rotation. The samples were subsequently purified with 300 µL chloroform:isoamyl-alcohol (24:1 BioUltra, Sigma-Aldrich). This involved mixing, then rotation at room temperature for 3 min, followed by centrifugation for 3 minutes at 3000 rpm. The supernatant was removed, then purified using MinElute spin columns (Qiagen, California, U.S.A.). Each purification was done by mixing 125 µL of supernatant with 5 volumes of PB buffer, followed by centrifugation through the column for 1 minute at 6000 rpm. The columns were washed with 750 µL PE buffer and centrifuged for 1 minute at 10 000 rpm, then dried by centrifugation for 3 minutes at maximum centrifuge speed. DNA was eluted in 105 µL 1× TET buffer after incubation at 37 °C for 15 minutes before elution. The whole process took around 1.5 hours per 10 samples.

2.2.3. Method G – G2 DNA/RNA enhancer with phenol-chloroform protocol

Each pellet was first added to G2 DNA/RNA enhancer (Ampliqon A/S, Denmark), after which the DNA was extracted following the method described in Zepeda-Mendoza et al. (2018). This involves a CTAB/PVP lysis buffer that comprised of 20 mM EDTA (0.5 M

UltraPure, pH 8.0, Invitrogen), 10 mM Tris-HCL (1M, pH 8.0), 1.4 M Sodium chloride (Sigma-Aldrich), 2% (w/v) Hexadecyltrimethylammonium bromide (CTAB, Sigma-Aldrich), 1% (v/v) 2-mercaptoethanol (99% BioUltra, Sigma-Aldrich), 2% (w/v) polyvinylpyrrolidone (PVP, Sigma-Aldrich), 0.05% Tween-20, 10% (v/v) Proteinase K (20 mg/mL, Recombinant PCR grade, Sigma-Aldrich) and filled up with Lonza AccuGene molecular biology water. DNA was eluted in 105 µL 1× TET buffer after incubation at 37 °C for 15 minutes before elution. The process took around 2.5 - 3 hours per 10 samples.

Table 2. Overview of extraction methods used in this study.

	Method F	Method C	Method G
Description (Manufacturer/ reference)	Commercial kit - FastDNA™ Spin Kit for Soil (MP Biomedical, (Piao et al., 2015; Stefanini et al., 2016)	CTAB/ PVPP with Chloroform (modified from (Rodríguez-Plaza et al., 2006)	G2 DNA/RNA enhancer with phenol-chloroform (Zepeda-Mendoza et al. 2018)
Homogenization	Bead-beating	Bead-beating	Only rotation with beads
pH adjustment	Yes	No	No
Thermal lysis	No	Yes	Yes
Chemical lysis	MT buffer (SDS, CTAB)	CTAB, 2-mercaptoethanol, proteinase K	Lysozyme, 2-mercaptoethanol, CTAB, proteinase K

Presence of hazardous chemicals	Binding Matrix (Guanidine Thiocyanate)	Chloroform:isoamyl-alcohol, 2-mercaptoethanol, Buffer PB (Guanidine hydrochloride)	Phenol-chloroform, 2-mercaptoethanol, Buffer PB (Guanidine hydrochloride)
Inhibitor(s) removal reagents	MT buffer (PVP, CTAB)	PVPP, CTAB, chloroform:isoamyl-alcohol	G2 DNA/RNA enhancer, PVP, CTAB, phenol-chloroform
Purification	Binding Matrix (silica matrix) and Spin™modules	Silica spin column (Qiagen)	Silica spin column (Qiagen)
Extra purification after DNA precipitation	DNA Clean and Concentrator™-5 (Zymo Research)	No	No
Handling time (10 samples)	1 hour 15 minutes	1.5 hours	2.5-3 hours
Price per samples (USD)	~8	~6	~18
Ease of use	Medium: requires careful handling of spin columns with Binding Matrix	Easy: requires careful handling of zirconia/ silica beads	Medium-Difficult: requires careful preparation and handling of G2 enhancer tubes; and spin column if using vacuum manifold

2.2.4. Quantitative Polymerase Chain Reaction (qPCR)

Quantitative PCR (qPCR) was used to estimate the number of copies of target region

within each extract, identify the presence of inhibitors, and quantify the number of cycles needed for metabarcoding PCR. The primers targeting the internal transcribed spacer 2 region (ITS2 from (Ihrmark et al., 2012) and (White et al., 1990)) and D2 domain of the large subunit of 26S rRNA gene (D2 from (Putignani et al., 2008) and (O'donnell, 1993)) (Table S1A), with primer sets modified into fusion primers with addition of both unique 8 bp multiplex identifier tag (MID tag) and Illumina MiSeq sequencing adapters (Table S1B in Supplementary file). We elected to use the fusion primers as it enables sequencing of the PCR products without further rounds of PCR and/or library construction, and thus reduces the introduction of subsequent biases (Murray et al., 2015). Standard curves for quantification were generated through a 10-fold serial dilution of PCR products by each primer (10^1 - 10^9 copies/μL), and were included in duplicate in all qPCR runs. Positive and negative (without DNA template) controls were also included.

qPCR assays were performed independently for each primer set at the following annealing temperatures: ITS2 (54 °C) and D2 (52 °C). Each qPCR setup consisted of 25 μL reaction volumes containing 2 μL of template and 23 μL of mastermix containing 1× GeneAmp ® 10X PCR Buffer II (Applied Biosystems, U.S.A.), 2.5 mM $MgCl_2$ (Applied Biosystems), 1 μL SYBR Green (Invitrogen, CA, USA), 0.8 mg/ mL Bovine Serum Albumin (BSA), 1 μL SYBR Green (Invitrogen), 0.25 mM of each dNTP, 0.4 μM forward primer, 0.4 μM reverse primer, 0.25 μL AmpliTaq Gold DNA polymerase (Applied Biosystems) and 14.5 μL AccuGene molecular biology water (Lonza). qPCR conditions were 95 °C for 5 minutes, followed by 45 cycles of 95 °C for 30 seconds, annealing for 30 seconds (at prior mentioned temperatures), and elongation at 72 °C for 45 seconds. A dissociation curve was performed as 1 cycle of 95 °C for 1 minutes, 55 °C

for 30 seconds and 95 °C for 30 seconds. All qPCR was performed using a MX3005 qPCR machine (Agilent). A random subset of two-thirds of each set of extracts was also diluted both 10 and 20 fold prior to qPCR, to assess for inhibition.

2.3 Metabarcoding PCRs and Sequencing

Metabarcoding PCR was performed for both ITS2 and D2 primers. The PCR mastermix was the same as that in qPCR, except replacing 1 μL SYBR Green (Invitrogen, CA, USA) with 1 μL AccuGene molecular biology water (Lonza, Switzerland). For the ITS2 region, 33 PCR cycles were used, while for the D2 domain, 35 PCR cycles were used. PCRs were carried out in either AB 2720 (Applied Biosystems, U.S.A.) or Veriti 96 well Thermal cyclers (Applied Biosystems, U.S.A.) with the following conditions. Initial incubation at 95 °C for 5 minutes, followed by cycles of 95 °C for 30 seconds, annealing for 30 seconds (at aforementioned temperatures), and extension at 72 °C for 45 seconds. The final elongation step was performed 72 °C for 10 minutes. Positive controls, extraction blanks and PCR negative controls were included in each PCR run. PCR products (ITS2: ~480 bp; D2: ~370 bp) were visualized by electrophoresis on 2% agarose gels, and subsequently pooled into 4 final pools for each target region. Pools were purified with QiaQuick columns (Qiagen) following the manufacturer's protocol to remove primer dimers. An aliquot of each pool was used for quantification and size estimation using the High-Sensitivity D1000 Screen Tape for Agilent 2200 TapeStation (Agilent). Lastly, purified PCR product pools were sequenced across a total of 4 flowcells (2 per target region) on an Illumina MiSeq platform using 250 bp paired end chemistry at The Danish National High-Throughput DNA Sequencing Centre. The sequencing data were deposited to European Nucleotide Archive under study number: PRJEB29796.

2.4 Bioinformatics analyses

Metabarcoding data analyses were performed using the pipeline described in Feld et al. (2016), although with modifications in trimming, post-clustering and the use of databases. Raw reads were merged with the command *-fastq_allowmergestagger* and demultiplexed using *vsearch* v2.1.2 (Rognes et al., 2016). Adaptors and primers were removed using *cutadapt* v1.11(Martin, 2011). Reads smaller than 100bp (ITS2) and 200 bp (D2) were trimmed with *vsearch*, followed by dereplication. Using the UPASE pipeline (Edgar, 2013), the dereplicated reads were subsequently filtered for singletons and chimeras, and clustered to operational taxonomic units (OTUs) with the command *-cluster_otus*. Reads were mapped back (including singletons) to the filtered clusters with 99% similarity using *usearch* v9.0.2132 (Edgar, 2010) in order to create the OTU table. The R-package LULU (Frøslev et al., 2017), as implemented in R v3.4.1, was used to improve the reliability of the OTU diversity estimates (Frøslev et al., 2017), by removing OTUs sequences based on similarity and co-occurrence patterns. Filtered OTUs derived from the ITS2 primers were then aligned with the reference UNITE+INSD database, released on 2017.12.01 (UNITE Community 2017), for taxonomic assignment at the genus level, with 97% identity threshold, 70% coverage in BLAST using QIIME v1.9.1 (Caporaso et al., 2010) with a modified *assign_taxonomy.py* script. OTUs that were not assigned to a taxon in the above were labelled as “No blast hit”. A similar approach was used to identify the OTUs generated using the D2 primers, with assignment performed against a curated large subunit database.

2.5 Statistical analysis

All statistical analyses were done using R v3.4.0, with Venn diagrams produced using Python v3.6.3. All the figures were visualized using the Matplotlib v3.0.0 (Hunter, 2007) and Seaborn v0.9.0 (Waskom et al., 2017) libraries with further post-processing in Inkscape v0.91 (https://inkscape.org/).

The DNA concentrations (Qubit value) of each sample and the qPCR results (log_{10} transformed number of copies of each marker, and threshold cycles (C_t) values) were individually analysed using Kruskal-Wallis rank sum test to test for differences between the three extraction methods, if significant, followed by Dunn test (Dunn, 1964; Ogle et al., 2018) for multiple comparisons using the Benjamini-Hochberg method for controlling FDR at 5% (Benjamini and Hochberg, 1995).

OTUs containing fewer than 10 reads, and samples with fewer than 1000 reads in total were discarded from further analyses (Oliver et al., 2015; Werner et al., 2012). Furthermore, OTUs that could not be taxonomically identified to genus (i.e. labeled as "No Blast Hit") were also removed from analyses. Venn diagrams were generated to represent the difference in presence/absence of OTUs between extraction methods at family level. The Venn diagram visualization script was modified from https://github.com/tctianchi/pyvenn. The family level network plot was visualized with Cytoscape (v. 3.7.0.).

Most downstream analyses were performed using the R package *phyloseq (McMurdie and Holmes, 2013)* unless noted otherwise. The α-diversity, the diversity of each sample, was estimated by rarefying samples to the lowest number of total reads observed from each sample. The α-diversity indices were calculated using the following metrics: Observed OTUs, Chao1, Simpson and Shannon. Pairwise α-diversity differences between the DNA extraction methods were tested using paired Wilcoxon signed-rank test (Wilcoxon, 1945) and for the pairwise extraction method and vineyard comparisons, the familywise error rate was adjusted using Bonferroni correction (familywise error rate = 0.05).

The β-diversity, the difference of between samples, was estimated using the following metrics: Bray-Curtis dissimilarity index, Jaccard index and Poisson dissimilarity. Bray-Curtis dissimilarity index was calculated using the *vegan* package in R (Oksanen et al., 2018). The original counts used for the input distance matrix were subjected first to variance stabilizing transformation with fittype 'mean' (Love et al., 2014) followed by min-max scaling instead of rarefying (McMurdie and Holmes, 2014). For the Jaccard index, the OTU table abundances were converted into presence/absence indicators. Poisson dissimilarity matrix was calculated from the original OTUs counts (Witten, 2011) with the *PoiClaClu* package. All matrices were subsequently decomposed for exploratory analysis using principal coordinates analysis (PCoA). The significance of differences among extraction methods and vineyards was evaluated with permutational multivariate analysis of variance (PERMANOVA) (Anderson, 2001), using distance matrices with 999 permutations using the *adonis* function, while homogeneity of group dispersions (PERMDISP2) (Anderson, 2006) was evaluated by applying the *betadisper* function from the *vegan* package (Oksanen et al., 2018). In order to estimate the significance of

PERMDISP2, classical analysis of variance (ANOVA) of the distances was performed together with Tukey's range test as a post-hoc test for multiple comparisons of means, at 5% family-wise error rate.

Differential abundance analysis was performed using the *DESeq2* R package (Love et al., 2014), in order to investigate the effects of different extraction methods on abundance of individual OTUs. *DESeq2* function parameters were set as following: test type: 'Wald'; fittype: 'parametric'. Significant differences between relative abundances were controlled for by setting FDR at 10% (Benjamini and Hochberg, 1995). The complete statistical analysis pipeline is available at https://github.com/kkpsiren/extmet. All the figures were visualized using the Matplotlib (Hunter, 2007) and Seaborn (Waskom et al., 2017) libraries with further post-processing in Inkscape v0.91 (https://inkscape.org/).

3. Results

3.1 Quantitative PCR (qPCR) results and DNA quantification

The concentration of extracted DNA varied among different extraction methods and samples (Fig. S1A). Samples extracted by methods C (1.54±1.87 ng/µl) and G (1.55±1.66 ng/µl) gave significantly higher concentrations than method F (0.28±0.132 ng/µl) (Table S2 and Fig. S1A). Similar results were obtained at the level of individual vineyards - method F always had a significantly lower DNA concentration than the other two methods (Fig. S1A and Table S2).

For the number of copies, no significant differences in ITS2 were observed (Table S2). Details of qPCR results can be found in Table S2. For the D2 marker, method F had significant lower number of copies obtained compared to method C (p<0.001) and G (p<0.001) (Fig. S1A and Table S3). Similar results were observed with the number of D2 copies among vineyards - method F had a significant influence compared to other 2 extraction methods (Table S3).

3.2 Metabarcoding Sequencing

3.2.1 ITS2

In total, 12,856,016 raw reads were obtained for the ITS2 marker, of which 12,737,449 were retained after removing adapters and primers. After all quality filtering and removal of reads from positive and negative controls, a total 12,012,768 reads were retained. One sample (w4b112) had fewer than 1000 reads, and was discarded. In total, 113 OTUs were retained for further analysis (details in Supplementary file). These included only 29 OTUs that were commonly shared between the 3 extraction methods (Fig. 1A). Method F yielded the highest number, with 72 total and 35 unique OTUs, while method C had the lowest number, with 43 total and 5 unique OTUs (Fig 1A). Overall these OTUs represent 49 fungal genera spanning 40 families and 25 orders (Fig. 2 and Table S4). Of these, the *Saccharomyces* genus was dominant, representing ~79% of the total reads counts, followed by *Starmerella*, class Dothideomycetes and *Metschnikowia* (Fig. S2A). In seventeen samples, the relative abundance of *Saccharomyces* was higher than 99% across every extraction method (Fig. S2A). The most diverse fungal community was observed in samples from vineyard 2b (F: 34; C: 18; G: 35 OTUs), vineyard 4b (F: 25; C: 24; G: 40 OTUs), and first sampling date of vineyard 5 (F: 21; C: 23; G: 34 OTUs) in all extraction

methods (Fig. S2A and Table S4).

3.2.2 D2

For the D2 target region, 18,246,033 raw reads were generated, with 14,982,993 reads retained following the adapter and primer removal. After quality filtering and removal of reads found in positive and negative controls, 14,085,938 reads remained for further analysis (details of reads filtering in Supplementary file). Although the total number of sequencing reads retained in D2 was larger than that in ITS2, 11 samples (w2a109, w2a120, w2a220, w2b214, w2b314, w3b220, w4a309, w4a312, w4b112, w4b312 and w4b120) contained fewer than 1000 reads; thus were discarded from further analysis (Fig. S3B). Details of the total number of reads per sample returned by the different extraction methods for both ITS2 and D2 are shown in Fig. S3 and Table S4.

Only 25 OTUs were retained after all filtering steps, of which 16 OTUs were found to be shared between the 3 methods (Fig 1B). As with the ITS2 results, method F had the highest total number (24) of OTUs and the highest number (4) of unique OTUs; however, both methods C and G did not have any unique OTUs (Fig 1B). These OTUs originated from 16 genera, and the samples were dominated by 2 genera, namely, *Hanseniaspora* (57.9%) and *Saccharomyces* (42%), with regards to the relative proportions of total read counts (Fig. S2B).

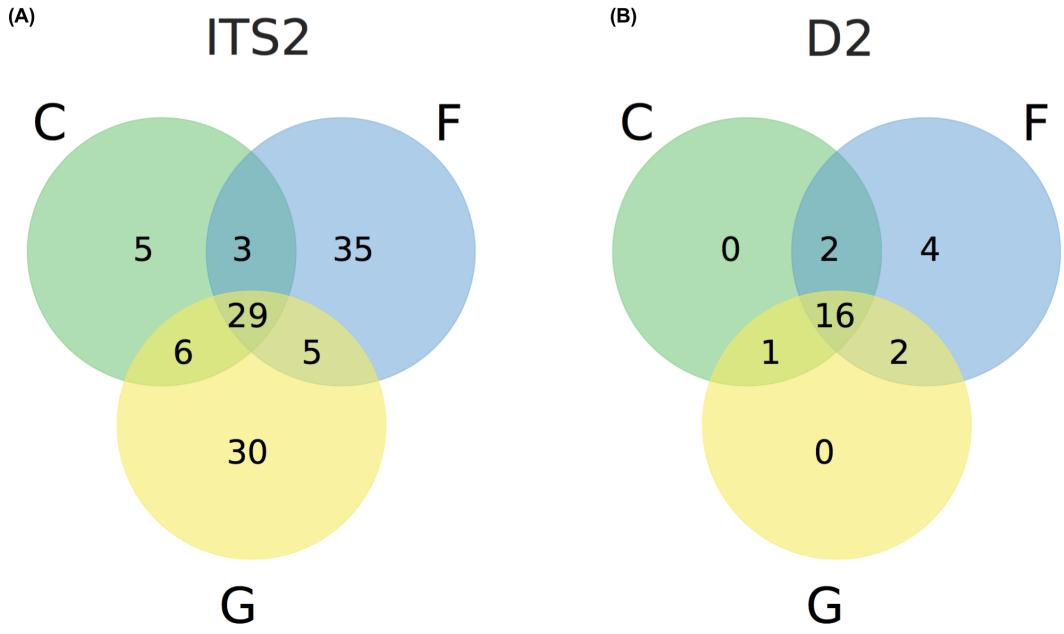


Fig 1. The number of OTUs obtained and shared by different extraction methods: C (green), F (blue) and G (yellow) in different genetic markers: (A) ITS2 (*n*=95) and (B) D2 (*n*=85).

To investigate the relationship between the ITS2 and D2 markers, the identification in two levels, family and genus were compared. Thirteen fungal families were found to be shared between the two markers, representing some of the most common wine related fungal families, such as Aspergillaceae, Cladosporiaceae, Metschnikowiaceae, Saccharomycetaceae and Saccharomycodaceae (Fig. 2). Additional taxonomic differences between the two marker regions were also observed. For instance, some of the most common non-*Saccharomyces* yeast: *Candida* was only observed in ITS2 but not in D2 marker while *Pichia* was found only in D2, not in ITS2 (Table S4).

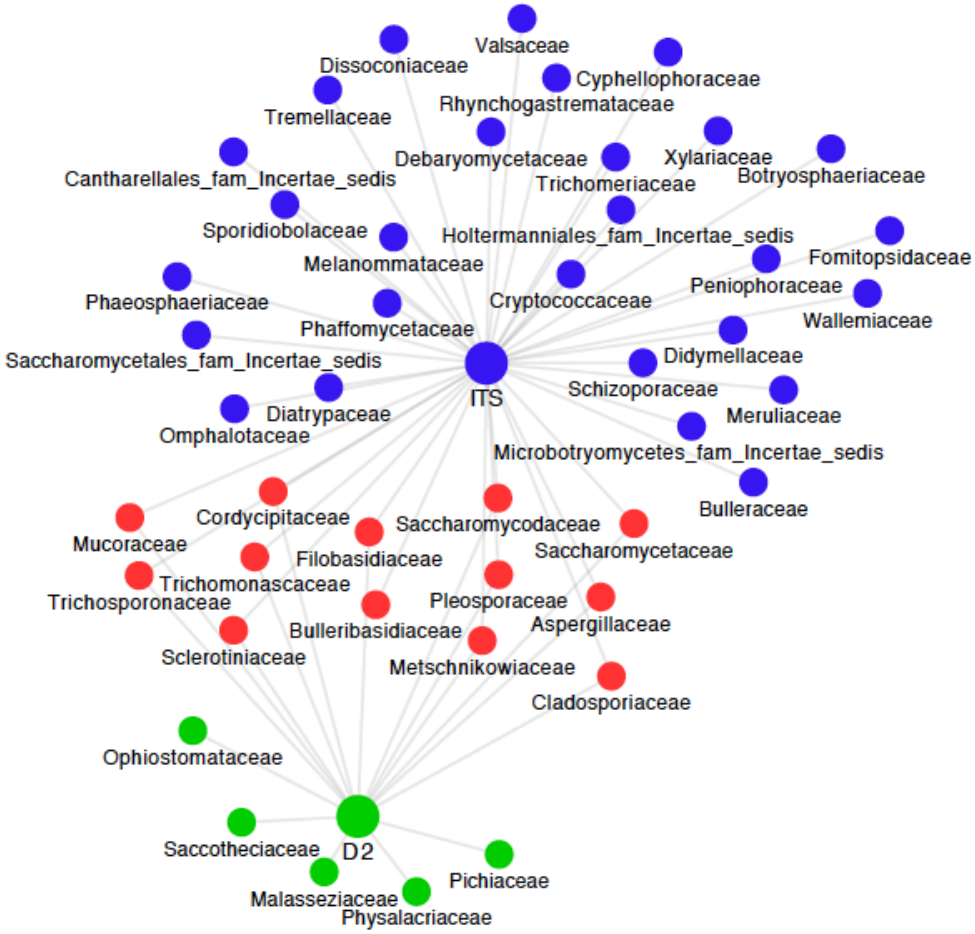


Fig. 2 Fungal families shared (red) and observed in both ITS2 (blue) and D2 (green) markers.

3.3 Species richness and community similarity

3.3.1 α-diversity

For the ITS2 region, pairwise comparison of the α-diversity revealed that the samples treated with extraction method G had a slightly, though significantly higher species richness measured with observed fungal OTUs, than that found in using the other methods (method C: $p<0.001$; method F: $p=0.003$). Similarly, significant differences in α-diversity were found between method G and the other two methods when estimated using the Chao1 index (method F: $p=0.01$; method C: $p<0.001$). Additionally, as the samples coming from the vineyards showed considerable variance (Fig.S5A), the effect of extraction methods on individual vineyards were also investigated. When investigating individual vineyards, species richness in method G was significantly higher for vineyard

2 than method C (Observed, C: p=0.014; Chao1, C: p=0.008) (Table S3D). No additional significant differences were observed between the extraction methods.

Fungal diversity among the extraction methods was also investigated using Shannon and Simpson diversity indices (Fig. S1B). Shannon and Simpson indices were found to be significantly lower in method C compared to methods F (Shannon: p=0.005; Simpson: p=0.009) and G (Shannon: p=0.002; Simpson: p=0.005) (Table S3C). Nonetheless, it was observed that for individual vineyards these indices were not affected by different extraction methods, apart from vineyard 2 where method C yielded significantly lower species diversity than method G (Shannon & Simpson: p=0.008).

For the D2 marker, pairwise comparison of the α-diversity revealed that the species richness of samples used in extraction method F was higher than that of method C (Chao1: p=0.015) (Table S3C). However, after Bonferroni correction for multiple comparisons, no further differences were observed. In general, the α-diversity, when measured with observed OTUs, was much lower for the D2 than ITS2 marker (Fig. S1B). Investigating the Shannon diversity and Simpson dominance indices, method F was found to be significantly lower than method C (Shannon: p=0.005 & Simpson:p=0.003; to method G, Shannon: p=0.026 & Simpson: p=0.022) (Table S3C). The effect of extraction methods on individual vineyards was also investigated, but no significant differences were observed (Table S3D).

3.3.2 β-diversity

The community structure was investigated using classical metrics, Bray-Curtis dissimilarity and Jaccard index, and additionally Poisson dissimilarity. For exploratory analysis, these metrics were decomposed by PCoA in both ITS and D2 (Fig. 3 and Fig. S4). The effect of extraction method was most pronounced with the Poisson dissimilarity in ITS2 (Fig. 3E), whereas all other metrics highlighted the vineyard effect over the extraction method (Fig. 3). PERMANOVA was used to partition the distance matrices among sources of variation originating from different extraction methods and vineyards. For all β-diversity metrics for both ITS2 and D2, we observed that both the vineyards and extraction methods were significantly different. Additionally, the interaction of the two variables was also found significant for ITS2 (Table S5 adonis). Furthermore, by investigating the coefficient of determination (R^2) and the pseudo-F (F-values) values, the vineyards seemed always to have a stronger effect than any extraction methods (Table S5). This was true in all cases except with Poisson dissimilarity in ITS2, where the R^2 was similar, and the pseudo-F value was higher for extraction method (Vineyards: F-value= 10.09, R^2= 0.186; Extraction methods: F-value= 15.05, R^2= 0.185). Investigating corresponding variable loadings, the OTUs unique to method F seemed to impact the distribution of the Bray-Curtis and Poisson dissimilarities (Fig. 3).

Moreover, by running pairwise PERMANOVAs among the extraction methods in ITS2, it was found that method F always differed from the other methods regardless of the metric, whereas the vineyard effect always was found higher despite the pairwise extraction comparisons (Table S5). The vineyard effect also always dominated in D2, while the extraction methods had less impact (Table S5).

3.3.3 Dispersion analysis

In order to investigate multivariate dispersion of the different extraction methods, the average distances of samples to the spatial medians of extraction methods were calculated using all distance metrics, from both markers individually. The PERMDISP2 results found significance in dispersion among extraction methods, for all the metrics except the ITS2 Poisson dissimilarity (Table S5). Bray-Curtis dissimilarity seemed to have the strongest impact on the variance (D2: F-value=6.49, p=0.0024, ITS2: F-value=4.94, p=0.0092). Pairwise comparisons revealed that the significant differences were always between method F and the other two. No significant differences were observed among methods G and C (Table S5).

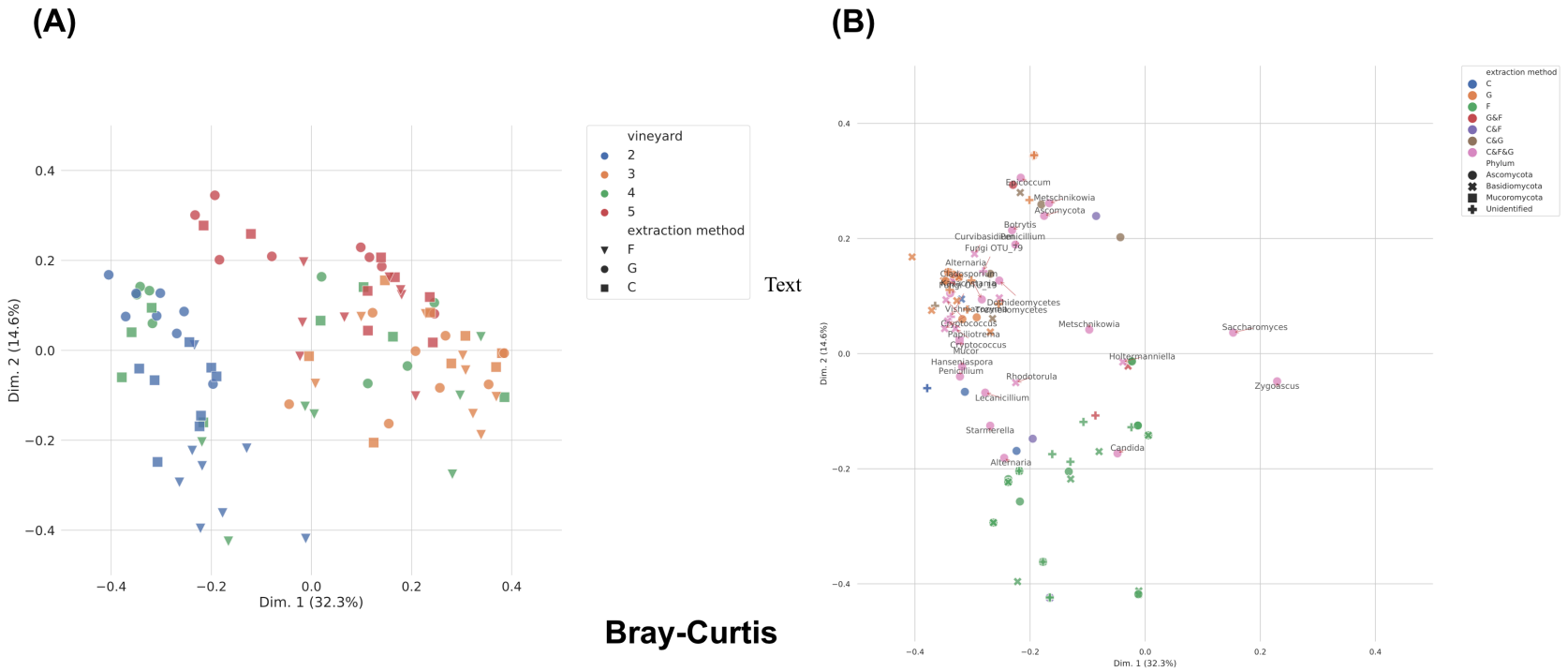


Fig. 3A & B β-diversity of fungal communities (left) and the corresponding variable loadings (right) from ITS2. PCoA plots from β-diversity of Bray-Curtis dissimilarity separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

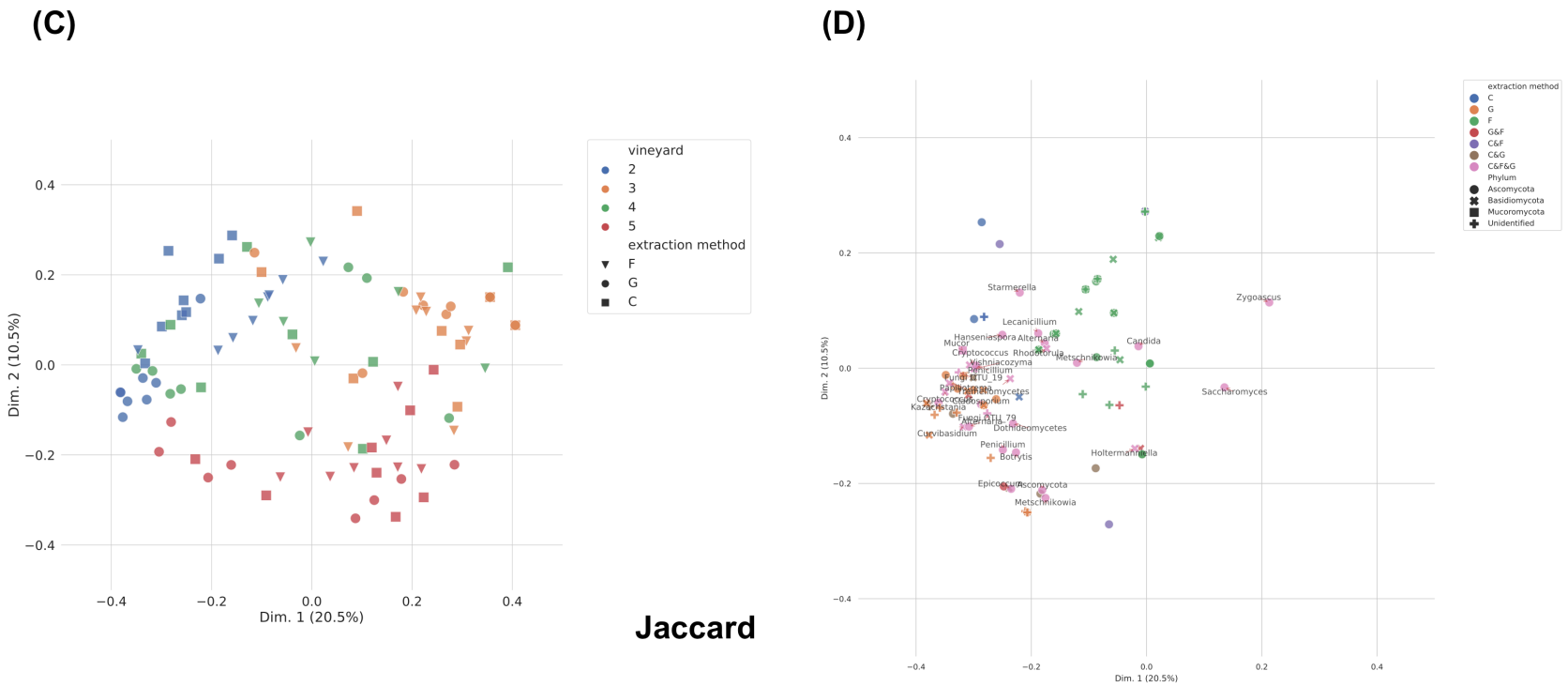


Fig. 3C & D β-diversity of fungal communities (left) and the corresponding variable loadings (right) from ITS2. PCoA plots from β-diversity of Jaccard index separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

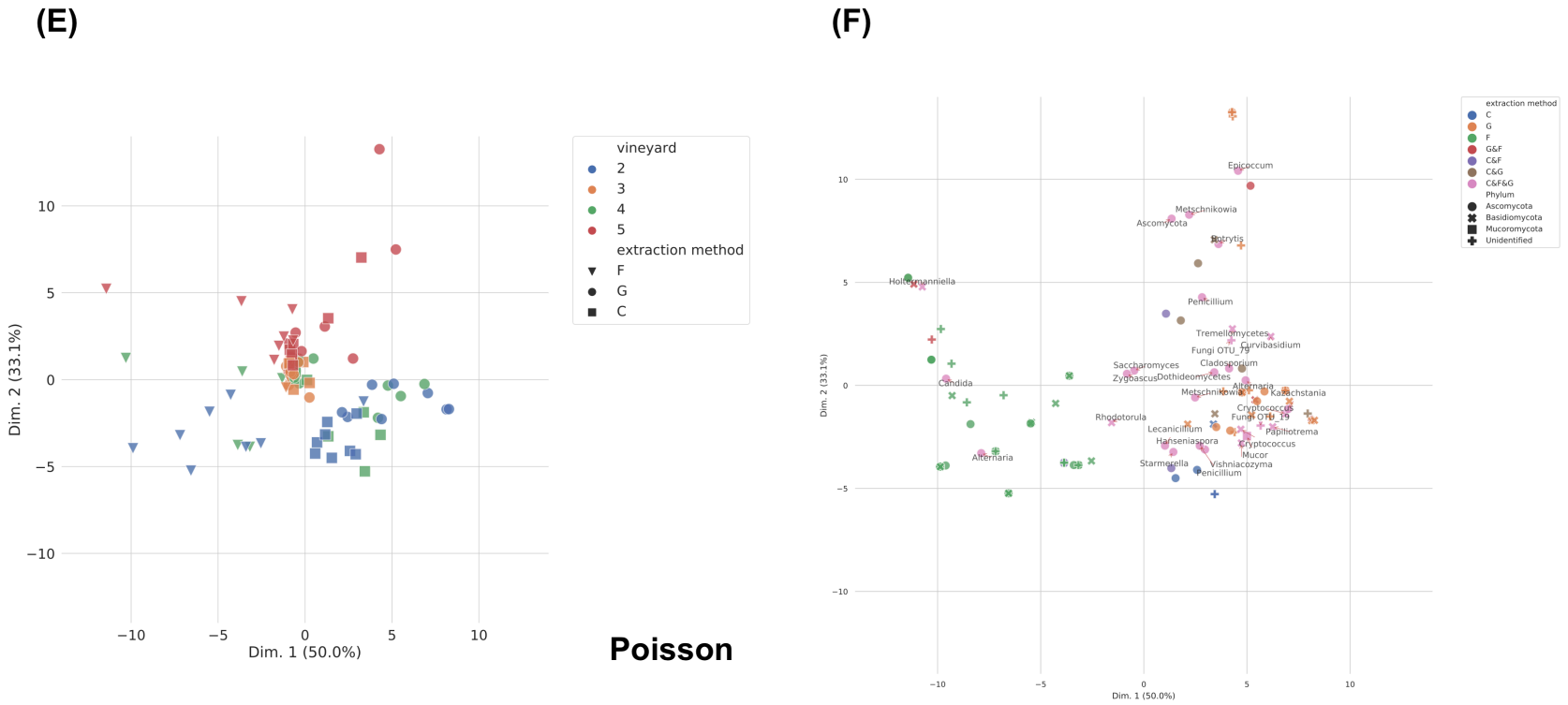


Fig. 3E & F β-diversity of fungal communities (left) and the corresponding variable loadings (right) from ITS2. PCoA plots from β-diversity of Poisson dissimilarity separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

3.4 Differential abundance

For D2, the pairwise differential abundance analysis revealed that two OTUs were significantly higher in method C than G (Table S6). These two OTUs were annotated to *Pichia* (*p-adj*=0.0018) and *Hanseniaspora* (*p-adj*=0.0022) (Fig. 4). No other significant differences among the extraction methods were observed while controlling of false positive discoveries.

For ITS2, pairwise differential abundance analysis revealed that 10 OTUs were significantly different between methods G and F, while 6 OTUs were significantly different between methods C and F. These 6 OTUs belonged to the above 10 OTUs, of which *Cryptococcus* was found twice in the comparison between methods G and F (Table S6). No significant differences were observed between methods G and C. Though only one OTU was significantly higher in method C, its taxonomy could not be resolved, and therefore was annotated as Fungi (Fig. 4).

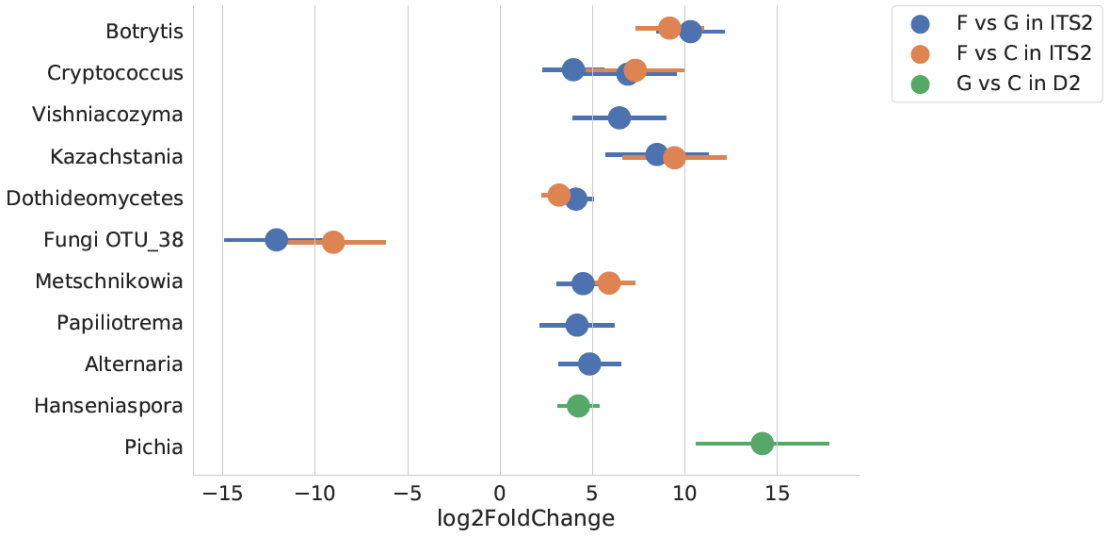


Fig. 4. The fold change (with log 2 standard errors) of significant OTUs differential abundance of 2 markers between different extraction methods. Two OTUs (green) were significantly different in comparing methods G and C in D2 marker. Ten OTUs were significantly different in comparing methods F and G (blue) while 6 OTUs were significantly different between methods F and C (orange) in ITS2 marker.

3.5 Ease of use, handling time and cost

Besides the fungal diversity recovered, often the decision on which particular extraction method to adopt relates to its ease of use, handing time, and cost (Table 2). Of the three methods, although Method C was only the second fastest method, it was the easiest to use. This is due to having only a single lysis step with bead-beating, single chemical purification step and a spin column purification step before DNA elution; however, the 0.1mm zirconia/silica beads used in bead-beating required care to prevent the beads splashing out.

Method F was moderately easy to use, and the fastest overall, as there was no thermal lysis step compared to methods C and G. However, care is required when handling the purification spin column and adding the kit's Binding Matrix, as the column lid is attached to the collection tube, which requires users to open the lid and remove the filter in order to discard the throw-through. Also, caution is needed when mixing in the Binding Matrix to prevent splashing.

The usability of Method G was moderate to difficult, and it was also the most time consuming - its protocol consisted of two 1-hour thermal lysis steps. Preparation of G2 DNA/RNA enhancers and lysis buffers required a substantial amount of time. Extra attention was needed in controlling the suction power of vacuum manifold. In all methods, the handling time would be extended with the increase in the number of samples.

The prices were obtained from either the manufacturer's websites or quotation from manufacturers. Method C was found to be the most economical while method G was the most expensive per sample (Table 2). Additionally, the MinElute spin columns used in Methods C and G would be substituted by magnetic silica beads which can further reduce the prices.

4. Discussion

4.1 General performances of extraction methods

In this study, we extracted DNA from must samples during alcoholic fermentation using three different DNA extraction methods, and undertook metabarcoding using two different genetic marker regions in order to investigate fungal diversity. We found that while method F always yielded the lowest DNA concentration, it yielded the largest number of OTUs (Fig. 1). This could be because different DNA extraction protocols can lead to highly variable total DNA yields, but this is not necessarily correlated to the targeted PCR-amplifiable microbe DNA (Josefsen et al., 2015) and also the PCR efficiency. Methods C and G had a similar assignment of OTUs, although method C was most conservative, yielding fewer OTUs in total.

Throughout all extraction methods, CTAB, and PVP or PVPP, were the common reagents found in the lysis buffers. CTAB is known to effectively remove polysaccharides from DNA extractions, and acts as a cationic surfactant to distort the cell membranes (Lever et al., 2015; Tan and Yiap, 2009). PVP and PVPP are both used for removing phenolic compounds and polysaccharides (Lever et al., 2015; Pereira et al., 2011), although a

previous study reported that PVPP gave a lower yield than PVP when extracting DNA from leaves due to insoluble cross-linking properties (Porebski et al., 1997). Moreover, additional chemical purifications applied chloroform-isoamyl alcohol and phenol-chloroform-isoamyl alcohol respectively in methods C and G, that could effectively remove the proteins, lipids and detergents (Lever et al., 2015), could enhance the DNA yield. The incorporation of G2 DNA/RNA enhancers in method G was aimed at improving the DNA extraction efficiency (Jacobsen et al., 2018). This could explain the higher DNA concentrations obtained in methods C and G.

Additionally, it has to be noted that standardisation makes experiments more comparable and reproducible. While "handcrafted" extraction protocols are often more economical and could give a higher DNA yield, the commercial kits are popular due to the streamlined procedures - not least because of ready-to-use reagents and applicability for various sample types. Despite the advantages of using commercial kits, some studies have identified how bacterial contamination can occur when using commercial kits (Glassing et al., 2016; Salter et al., 2014). However, so far no studies have investigated the fungal contamination from kits or PCR reagents. From our results, each extraction negative control from method F (commercial kit) did not show any definitive or observable contamination. It will be useful to conduct some studies of the potential fungal contamination for commercial extraction kits in the future.

4.2 α-diversity effects

The within sample diversity was found to vary as expected, with the main source of variation originating from the vineyards (referred as vineyard effect in the following

texts) (Fig. S1B). No clear conclusions could be said from the performance α-diversity measurements but method C was observed to have the lowest species richness within sample in both ITS2 and D2. However, in ITS2, method G had the highest amount of observed OTUs, mainly because it was able to extract information out of samples from vineyard 2 that were observed to struggle in alcoholic fermentation (data not shown).

4.3 Interpretation of the β-diversity effects

PERMANOVA has been shown in simulation studies to be the preferred analysis especially when heterogenous dispersions are observed in data (Anderson and Walsh, 2013). While the analysis still assumes homogeneity of within group variances, it is established that a robust test for homogeneity of multivariate dispersion is required (Anderson, 2006). This study included three different β-diversity metrics, as pattern recognition differences occur with different metrics (Kuczynski et al., 2010), but they can also have an impact on the observed dispersions (Anderson, 2006). Besides the choice of metric, particular attention has to be paid to how data is transformed (McMurdie and Holmes, 2014) as HTS data distribution tends to be highly positively skewed. Thus, Jaccard presence-absence, variance stabilized transformed (Love et al., 2014) Bray-Curtis and Poisson distance based (Witten, 2011) metrics were used to investigate extraction method and vineyard effects by mean changes and dispersion.

Interestingly, we observed that the variance within extraction methods was significantly different, no matter the marker region or dissimilarity measure used, with the sole exception of Poisson dissimilarity in ITS2. Further investigation of the multiple pairwise

comparisons revealed that extraction methods yielded no significant differences between methods G and C, whereas method F was almost always different (Table S5). Furthermore, in investigating the pairwise PERMANOVAs, the vineyard effect was always the most prominent between methods G and C in Bray-Curtis dissimilarities in both methods (Table. S5). This suggested that the effect of extraction method had less impact between these two extraction methods. For both markers, Jaccard distances were observed with less extraction method effect (Table S5). Therefore, most of the extraction method effect can be hypothesised to be originating from the abundance differences. This could relate to the overall level of fungal cells extracted as indicated with the higher DNA concentrations and copy numbers obtained in methods C and G. However, according to (Josefsen et al., 2015), the highly variable total DNA yields might not necessarily correlated to the targeted PCR-amplifiable microbe DNA as we still extract similar information (and more species) with method F.

Additionally, we observed that for both Jaccard and Bray-Curtis metrics, the dispersion was significantly affected by the extraction method. Although the biological explanation for Jaccard distance is less clear, it has been suggested that the increase in dispersion could relate to species loss in Bray-Curtis dissimilarity (Warwick and Clarke, 1993). Indeed, this was observed when looking into the relative differential abundance of the OTUs (Fig. 4) in ITS2. Method F was found to have significantly lower expression of 10 OTUs, some of which are related to winemaking (*Kazachstania*, *Metschnikowia* and *Botrytis*), while only one unidentified Fungi OTU was observed in higher abundance (Fig. 4). Nevertheless, only 4 genera among these 10 OTUs were well expressed when considering the mean of normalised sample counts (baseMean > 40, Table S6).

Additionally, while the Poisson dissimilarity homogenised the multivariate dispersion, it was shown that both extraction methods and vineyards had a significant effect on β-diversity measure. Furthermore, it could be noted from the pseudo-F values that method F had higher impact than vineyards, even though the proportion of variance explained (R^2 values) was similar (Table S5). Investigating the OTU contributions, it became clear that read count transformation to Poisson dissimilarities enhanced the OTU abundance differences, similar to differential abundance analysis. Therefore many unique OTUs in method F played a bigger role. Additionally, it was observed that the vineyard contributed to the separation of the second dimension in the ordination plot (Fig. 3E), which suggested that no matter which extraction methods and analyses used, vineyard effect was always retained. Nonetheless, when applying the Poisson dissimilarity, we can investigate the influence of differentially expressed OTUs (Witten, 2011) in extraction methods. Therefore we conclude that the main effect on the diversity of the fungal community was the vineyard instead of the chosen extraction method.

4.4 Fungal taxonomic diversities from ITS2 and D2 markers

According to the fungal taxonomic diversity obtained using the ITS2 marker, *Saccharomyces* always dominated in the number of reads in a majority of samples and extraction methods which is expected during wine alcoholic fermentation. During the alcoholic fermentation, the microbial diversity reduces with the increase of alcoholic percentage and alcohol tolerant *Saccharomyces* start to dominate the fermentation. In addition, other dominant winemaking related taxa like *Starmerella*, Dothideomycetes and *Metschnikowia* were mainly found in the early stage of fermentation samples, in line with

results from other studies (Bokulich et al., 2014; Pinto et al., 2015).

Twenty-five OTUs were observed with the D2 markers among all extraction methods. Apart from the 2 dominant genera *Hanseniaspora* and *Saccharomyces*, other winemaking related taxa such as *Pichia, Lachancea* and family Metschnikowiaceae were also found. Our findings are consistent with results from a study on fungal diversity in grape juice and ferments in New Zealand (Morrison-Whittle and Goddard, 2018), such as the presence of *Alternaria, Botrytis, Cladosporium, Hanseniaspora, Pichia, Lachancea* and *Saccharomyces*. Another genus of interest is *Torulaspora,* which was originally found in both studies; however, after running LULU with strict settings, it was found to be merged with *Saccharomyces* which is phylogenetically similar (Masneuf-Pomarede et al., 2015).

When comparing the fungal diversity between two marker regions, differences in taxonomy were observed. ITS2 markers obtained more OTUs across different taxonomic levels than D2 markers. Other previous studies also report more OTUs obtained for the ITS2 than D2 markers (Pinto et al., 2014; Stefanini and Cavalieri, 2018) Nonetheless, this contradicted a study about fermentation of Portuguese red wine grapes (Pinto et al., 2015), which observed slightly higher numbers of OTUs in D2 than ITS2. Our findings on the high abundance of *Hanseniaspora* in D2 compared in ITS2 (Fig. S2) might be due to the primer bias of 26S rRNA gene in amplifying *Hanseniaspora* genus (Mota-Gutierrez et al., 2018). Furthermore, other genera have been found to be underestimated (e.g. *Saccharomyces* in 26S), while others overestimated (eg. *Penicillium* in ITS2) by different marker regions (Mota-Gutierrez et al., 2018). Nevertheless, it is hard to say which region preferable in characterising fungal diversity as there are studies that use

both (Scorzetti et al., 2002) or solely 26S (Mota-Gutierrez et al., 2018) or ITS2 (Raja et al., 2017). In light of the pitfalls of choosing the marker regions, it would be a better idea to apply shotgun metagenomics for more in-depth approach if possible (Sirén et al., 2019).

4.5 Complexity of nomenclature in fungal databases

When comparing the fungal diversity difference between two marker regions at genus level in this study, it was found that some of the most common non-*Saccharomyces* winemaking related yeasts were only found in one of the markers. For instance, *Candida* was only observed in ITS2 but not in D2 marker while *Pichia* was found only in D2, not in ITS2 (Table S4). However, this dissimilarity could be related to the complexity of different morphs and synonyms of yeasts. The nomenclature of yeasts is complicated due to historical different naming of anamorphs and teleomorphs. With the advancement of molecular technology, the classifications of fungi (Tedersoo et al., 2018), especially the main wine related phylum Ascomycota (Wijayawardene et al., 2018) has seen subsequent changes and updates. According to Masneuf-Pomarede et al. (2015), the majority of *Pichia* species were originally classified to *Candida,* while some of the *Starmerella* species were also classified as *Candida*. The broad definition of *Candida* has lead to it becoming the dumping ground for phylogenetic placement (Kurtzman et al., 2015; Suh et al., 2006). Since the public available databases e.g. UNITE (Nilsson et al., 2018) and GenBank (Clark et al., 2016) act as the fundamental references for researches, if these changes do not get updated or validated in these databases, the on-going changes in nomenclature of yeasts will lead to errors in taxonomic assignments.

5. Conclusion

Vineyard effect on fungal diversity was observed in all extraction methods in both markers. Any appropriate extraction method should be considered individually based on sample types and scale, research budget, handling time and research questions since no universally standard, generically “optimal” extraction method exists. We have reviewed three extraction methods for standardising the wine microbial studies. Method G had the highest DNA yield, while the cost per sample of method F was in the middle among 3 protocols with the fewest amount of noticeable hazardous chemicals involved and somewhat shortest handling time. Method F also obtained the highest number of unique OTUs, so will be a useful method for studying fungal diversity in winemaking related datasets in general. Lastly, our results highlight the need to standardise DNA extraction methods if such data generated from wine microbe profiling studies are to be compared.

6. Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7. Acknowledgements

This work was funded by the Horizon 2020 Programme of the European Commission within the Marie Skłodowska-Curie Innovative Training Network “MicroWine” (grant number 643063). The authors would like to thank Lara Puetz and Shanlin Liu for bioinformatic advice, Engela Stadler in assisting the fermentation trial set up, and the collaborating wine estates in the Pfalz. Authors would also like to acknowledge the

Danish National High-Throughput Sequencing Centre for assistance in data generation.

8. Supplementary information

Supplementary File: Details of metabarcoding sequencing quality filtering
Supplementary Figures

Supplementary Tables

9. References

Anderson, M.J., 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62, 245–253. https://doi.org/10.1111/j.1541-0420.2005.00440.x

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance: non-parametric MANOVAa for ecology. Austral Ecol. 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson, M.J., Walsh, D.C.I., 2013. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing? Ecol. Monogr.

Belda, I., Navascués, E., Marquina, D., Santos, A., Calderón, F., Benito, S., 2016a. Outlining the influence of non-conventional yeasts in wine ageing over lees. Yeast 33, 329–338. https://doi.org/10.1002/yea.3165

Belda, I., Ruiz, J., Alastruey-Izquierdo, A., Navascués, E., Marquina, D., Santos, A.,

2016b. Unraveling the enzymatic basis of wine “flavorome”: a phylo-functional study of wine related yeast species. Front. Microbiol. 7, 243. https://doi.org/10.3389/fmicb.2016.00012

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. 57, 289–300.

Bisson, L.F., Joseph, C.M.L., Domizio, P., 2017. Yeasts, in: König, H., Unden, G., Fröhlich, J. (Eds.), Biology of microorganisms on grapes, in must and in wine. Springer International Publishing, Cham, pp. 65–101.

Bokulich, N.A., Ohta, M., Richardson, P.M., Mills, D.A., 2013. Monitoring Seasonal Changes in Winery-Resident Microbiota. PLoS One 8, e66437. https://doi.org/10.1371/journal.pone.0066437

Bokulich, N.A., Thorngate, J.H., Richardson, P.M., Mills, D.A., 2014. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. Proc. Natl. Acad. Sci. U. S. A. 111, E139–48. https://doi.org/10.1073/pnas.1317377110

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. https://doi.org/10.1038/nmeth.f.303

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2016. GenBank.

Nucleic Acids Res. 44, D67–72. https://doi.org/10.1093/nar/gkv1276

Cordero-Bueso, G., Rodríguez, M.E., Garrido, C., Cantoral, J.M., 2017. Rapid and not culture-dependent assay based on multiplex PCR-SSR analysis for monitoring inoculated yeast strains in industrial wine fermentations. Arch. Microbiol. 199, 135–143. https://doi.org/10.1007/s00203-016-1287-4

Dopheide, A., Xie, D., Buckley, T.R., Drummond, A.J., Newcomb, R.D., 2018. Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.13086

Dunn, O.J., 1964. Multiple Comparisons Using Rank Sums. Technometrics 6, 241–252. https://doi.org/10.2307/1266041

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998. https://doi.org/10.1038/nmeth.2604

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Feld, L., Nielsen, T.K., Hansen, L.H., Aamand, J., Albers, C.N., 2016. Establishment of bacterial herbicide degraders in a rapid sand filter for bioremediation of phenoxypropionate-polluted groundwater. Appl. Environ. Microbiol. 82, 878–887. https://doi.org/10.1128/AEM.02600-15

Frøslev, T.G., Kjøller, R., Bruun, H.H., Ejrnæs, R., Brunbjerg, A.K., Pietroni, C., Hansen, A.J., 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat. Commun. 8, 1188. https://doi.org/10.1038/s41467-017-01312-x

Glassing, A., Dowd, S.E., Galandiuk, S., Davis, B., Chiodini, R.J., 2016. Inherent

bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathog. 8, 24. https://doi.org/10.1186/s13099-016-0103-7

Godálová, Z., Kraková, L., Puškárová, A., Bučková, M., Kuchta, T., Piknová, Ľ., Pangallo, D., 2016. Bacterial consortia at different wine fermentation phases of two typical Central European grape varieties: Blaufränkisch (Frankovka modrá) and Grüner Veltliner (Veltlínske zelené). Int. J. Food Microbiol. 217, 110–116. https://doi.org/10.1016/j.ijfoodmicro.2015.10.015

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55

Ihrmark, K., Bödeker, I.T.M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandström-Durling, M., Clemmensen, K.E., Lindahl, B.D., 2012. New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol. Ecol. 82, 666–677. https://doi.org/10.1111/j.1574-6941.2012.01437.x

Inceoglu, O., Hoogwout, E.F., Hill, P., van Elsas, J.D., 2010. Effect of DNA extraction method on the apparent microbial diversity of soil. Appl. Environ. Microbiol. 76, 3378–3382. https://doi.org/10.1128/AEM.02715-09

Işçi, B., Kalkan Yildirim, H., Altindisli, A., 2014. Evaluation of methods for DNA extraction from must and wine. J. Inst. Brew. 120, 238–243. https://doi.org/10.1002/jib.129

Jacobsen, C.S., Nielsen, T.K., Vester, J.K., Stougaard, P., Nielsen, J.L., Voriskova, J., Winding, A., Baldrian, P., Liu, B., Frostegård, Å., Pedersen, D., Tveit, A.T., Svenning,

M.M., Tebbe, C.C., Øvreås, L., Jakobsen, P.B., Blazewicz, S.J., Hubablek, V., Bertilsson, S., Hansen, L.H., Cary, S.C., Holben, W.E., Ekelund, F., Bælum, J., 2018. Inter-laboratory testing of the effect of DNA blocking reagent G2 on DNA extraction from low-biomass clay samples. Sci. Rep. 8, 5711. https://doi.org/10.1038/s41598-018-24082-y

Jara, C., Mateo, E., Guillamón, J.M., Torija, M.J., Mas, A., 2008. Analysis of several methods for the extraction of high quality DNA from acetic acid bacteria in wine and vinegar for characterization by PCR-based methods. Int. J. Food Microbiol. 128, 336–341. https://doi.org/10.1016/j.ijfoodmicro.2008.09.008

Josefsen, M.H., Andersen, S.C., Christensen, J., Hoorfar, J., 2015. Microbial food safety: Potential of DNA extraction methods for use in diagnostic metagenomics. J. Microbiol. Methods 114, 30–34. https://doi.org/10.1016/j.mimet.2015.04.016

Keisam, S., Romi, W., Ahmed, G., Jeyaram, K., 2016. Quantifying the biases in metagenome mining for realistic assessment of microbial ecology of naturally fermented foods. Sci. Rep. 6, 34155. https://doi.org/10.1038/srep34155

Kuczynski, J., Liu, Z., Lozupone, C., McDonald, D., Fierer, N., Knight, R., 2010. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nat. Methods 7, 813–819. https://doi.org/10.1038/nmeth.1499

Kurtzman, C.P., Mateo, R.Q., Kolecka, A., Theelen, B., Robert, V., Boekhout, T., 2015. Advances in yeast systematics and phylogeny and their use as predictors of biotechnologically important metabolic pathways. FEMS Yeast Res. 15. https://doi.org/10.1093/femsyr/fov050

Lever, M.A., Torti, A., Eickenbusch, P., Michaud, A.B., Šantl-Temkiv, T., Jørgensen, B.B., 2015. A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. Front. Microbiol. 6, 476.

Lleixà, J., Kioroglou, D., Mas, A., Portillo, M.D.C., 2018. Microbiome dynamics during spontaneous fermentations of sound grapes in comparison with sour rot and Botrytis infected grapes. Int. J. Food Microbiol. 281, 36–46. https://doi.org/10.1016/j.ijfoodmicro.2018.05.016

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12. https://doi.org/10.14806/ej.17.1.200

Masneuf-Pomarede, I., Bely, M., Marullo, P., Albertin, W., 2015. The genetics of non-conventional wine yeasts: current knowledge and future challenges. Front. Microbiol. 6, 1563.

McMurdie, P.J., Holmes, S., 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput. Biol. 10, e1003531. https://doi.org/10.1371/journal.pcbi.1003531

McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8, e61217. https://doi.org/10.1371/journal.pone.0061217

Morrison-Whittle, P., Goddard, M.R., 2018. From vineyard to winery: a source map

of microbial diversity driving wine fermentation. Environ. Microbiol. 20, 75–84.

Mota-Gutierrez, J., Ferrocino, I., Rantsiou, K., Cocolin, L., 2018. Metataxonomic comparison between internal transcribed spacer and 26S ribosomal large subunit (LSU) rDNA gene. Int. J. Food Microbiol. https://doi.org/10.1016/j.ijfoodmicro.2018.10.010

Murray, D.C., Coghlan, M.L., Bunce, M., 2015. From benchtop to desktop: important considerations when designing amplicon sequencing workflows. PLoS One 10, e0124671. https://doi.org/10.1371/journal.pone.0124671

Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., Saar, I., Kõljalg, U., Abarenkov, K., 2018. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res. https://doi.org/10.1093/nar/gky1022

O'donnell, K., 1993. Fusarium and its near relatives, in: Reynolds, D.R. and Taylor, J.W. (Ed.), The fungal holomarph: mititic, meiotic and pleomorpic speciation in fungal systematic. ci.nii.ac.jp, pp. 225–233.

Ogle, D.H., Wheeler, P., Dinno, A., 2018. FSA: Fisheries Stock Analysis [WWW Document]. R package version 0.8.21.9000. URL https://github.com/droglenc/FSA

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2018. The vegan package version 2.5-2. Community ecology package. [http://r-forge. r-project. org/projects/vegan/].

Oliver, A.K., Brown, S.P., Callaham, M.A., Jumpponen, A., 2015. Polymerase matters: non-proofreading enzymes inflate fungal community richness estimates by

up to 15 %. Fungal Ecol. 15, 86–89. https://doi.org/10.1016/j.funeco.2015.03.003

Pereira, L., Guedes-Pinto, H., Martins-Lopes, P., 2011. an enhanced method for Vitis vinifera L. DNA extraction from wines. Am. J. Enol. Vitic. 62, 547–552. https://doi.org/10.5344/ajev.2011.10022

Piao, H., Hawley, E., Kopf, S., DeScenzo, R., Sealock, S., Henick-Kling, T., Hess, M., 2015. Insights into the bacterial community and its temporal succession during the fermentation of wine grapes. Front. Microbiol. 6, 809. https://doi.org/10.3389/fmicb.2015.00809

Pinto, C., Pinho, D., Cardoso, R., Custódio, V., Fernandes, J., Sousa, S., Pinheiro, M., Egas, C., Gomes, A.C., 2015. Wine fermentation microbiome: a landscape from different Portuguese wine appellations. Front. Microbiol. 6, 905. https://doi.org/10.3389/fmicb.2015.00905

Pinto, C., Pinho, D., Sousa, S., Pinheiro, M., Egas, C., Gomes, A.C., 2014. Unravelling the diversity of grapevine microbiome. PLoS One 9, e85622. https://doi.org/10.1371/journal.pone.0085622

Porebski, S., Bailey, L.G., Baum, B.R., 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Rep. 15, 8–15. https://doi.org/10.1007/BF02772108

Putignani, L., Paglia, M.G., Bordi, E., Nebuloso, E., Pucillo, L.P., Visca, P., 2008. Identification of clinically relevant yeast species by DNA sequence analysis of the D2 variable region of the 25-28S rRNA gene. Mycoses 51, 209–227. https://doi.org/10.1111/j.1439-0507.2007.01472.x

Raja, H.A., Miller, A.N., Pearce, C.J., Oberlies, N.H., 2017. fungal identification

using molecular tools: a primer for the natural products research community. J. Nat. Prod. 80, 756–770. https://doi.org/10.1021/acs.jnatprod.6b01085

Rodríguez-Plaza, P., González, R., Moreno-Arribas, M.V., Polo, M.C., Bravo, G., Martínez-Zapater, J.M., Martínez, M.C., Cifuentes, A., 2006. Combining microsatellite markers and capillary gel electrophoresis with laser-induced fluorescence to identify the grape (Vitis vinifera) variety of musts. Eur. Food Res. Technol. 223, 625–631. https://doi.org/10.1007/s00217-005-0244-2

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87. https://doi.org/10.1186/s12915-014-0087-z

Savazzini, F., Martinelli, L., 2006. DNA analysis in wines: Development of methods for enhanced extraction and real-time polymerase chain reaction quantification. Anal. Chim. Acta 563, 274–282. https://doi.org/10.1016/j.aca.2005.10.078

Scorzetti, G., Fell, J.W., Fonseca, A., Statzell-Tallman, A., 2002. Systematics of basidiomycetous yeasts: a comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. FEMS Yeast Res. 2, 495–517.

Sirén, K., Mak, S.S.T., Fischer, U., Hansen, L.H., Gilbert, M.T.P., 2019. Multi-omics and potential applications in wine production. Curr. Opin. Biotechnol. 56, 172–178. https://doi.org/10.1016/j.copbio.2018.11.014

Siret, R., Boursiquot, J.M., Merle, M.H., Cabanis, J.C., This, P., 2000. Toward the

authentication of varietal wines by the analysis of grape (Vitis vinifera L.) residual DNA in must and wine using microsatellite markers. J. Agric. Food Chem. 48, 5035–5040. https://doi.org/10.1021/jf991168a

Stefanini, I., Albanese, D., Cavazza, A., Franciosi, E., De Filippo, C., Donati, C., Cavalieri, D., 2016. Dynamic changes in microbiota and mycobiota during spontaneous “Vino Santo Trentino” fermentation. Microb. Biotechnol. 9, 195–208. https://doi.org/10.1111/1751-7915.12337

Stefanini, I., Cavalieri, D., 2018. Metagenomic approaches to investigate the contribution of the vineyard environment to the quality of wine fermentation: potentials and difficulties. Front. Microbiol. 9, 991.

Sternes, P.R., Lee, D., Kutyna, D.R., Borneman, A.R., 2017. A combined meta-barcoding and shotgun metagenomic analysis of spontaneous wine fermentation. Gigascience 6, 1–10.

Suh, S.-O., Blackwell, M., Kurtzman, C.P., Lachance, M.-A., 2006. Phylogenetics of Saccharomycetales, the ascomycete yeasts. Mycologia 98, 1006–1017. https://doi.org/10.1080/15572536.2006.11832629

Tan, S.C., Yiap, B.C., 2009. DNA, RNA, and protein extraction: the past and the present. J. Biomed. Biotechnol. 2009, 574398. https://doi.org/10.1155/2009/574398

Tedersoo, L., Sánchez-Ramírez, S., Kõljalg, U., Bahram, M., Döring, M., Schigel, D., May, T., Ryberg, M., Abarenkov, K., 2018. High-level classification of the Fungi and a tool for evolutionary ecological analyses. Fungal Divers. 90, 135–159. https://doi.org/10.1007/s13225-018-0401-0

Villano, C., Lisanti, M.T., Gambuti, A., Vecchio, R., Moio, L., Frusciante, L.,

Aversano, R., Carputo, D., 2017. Wine varietal authentication based on phenolics, volatiles and DNA markers: State of the art, perspectives and drawbacks. Food Control 80, 1–10. https://doi.org/10.1016/j.foodcont.2017.04.020

Warwick, R.M., Clarke, K.R., 1993. Increased variability as a symptom of stress in marine communities. J. Exp. Mar. Bio. Ecol. 172, 215–226. https://doi.org/10.1016/0022-0981(93)90098-9

Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lukauskas, S., Gemperline, D.C., Augspurger, T., Halchenko, Y., Cole, J.B., Warmenhoven, J., de Ruiter, J., Pye, C., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Bachant, P., Martin, M., Meyer, K., Miles, A., Ram, Y., Yarkoni, T., Williams, M.L., Evans, C., Fitzgerald, C., Brian, Fonnesbeck, C., Lee, A., Qalieh, A., 2017. mwaskom/seaborn: v0.8.1 (September 2017). https://doi.org/10.5281/zenodo.883859

Werner, J.J., Zhou, D., Caporaso, J.G., Knight, R., Angenent, L.T., 2012. Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J. 6, 1273–1276. https://doi.org/10.1038/ismej.2011.186

White, T.J., Bruns, T., Lee, S., Taylor, J.L., Others, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18, 315–322.

Wijayawardene, N.N., Hyde, K.D., Lumbsch, H.T., Liu, J.K., Maharachchikumbura, S.S.N., Ekanayaka, A.H., Tian, Q., Phookamsak, R., 2018. Outline of Ascomycota: 2017. Fungal Divers. 88, 167–263. https://doi.org/10.1007/s13225-018-0394-8

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics Bulletin 1, 80–83. https://doi.org/10.2307/3001968

Witten, D.M., 2011. Classification and clustering of sequencing data using a Poisson model. Ann. Appl. Stat. 5, 2493–2518. https://doi.org/10.1214/11-AOAS493

Zepeda-Mendoza, M.L., Edwards, N.K., Madsen, M.G., Abel-Kistrup, M., Puetz, L., Sicheritz-Ponten, T., Swiegers, J.H., 2018. Influence of Oenococcus oeni and Brettanomyces bruxellensis on wine microbial taxonomic and functional potential profile. Am. J. Enol. Vitic. ajev.2018.17092.

Supplementary information of Chapter 3

Supplementary information

Supplementary File 1 - Details of metabarcoding sequencing quality filtering

The fusion primers comprised of MiSeq sequencing adapters + MID index + target region primers. The MiSeq sequencing adapters used in the study were shown in Table S1B.

ITS2

In total, 12,856,016 ITS2 raw reads were obtained, of which 12,737,449 were retained after removing adapters and primers. Of these, 328,172 reads were retained and used to generate the OTUs (168 OTUs) after dereplication and removal of singletons. After running the LULU scripts, 158 OTUs were retained for further analysis. Finally, after removing all the reads from positive and negative control samples, a total 12,012,768 reads were retained. One sample (w4b112) had fewer than 1000 reads, and thus was discarded from further analysis. In total, 113 OTUs were retained after filtering, of which 29 OTUs were commonly shared by 3 extraction methods (Fig. 1A).

(a) D2

For D2 target region, 18,246,033 raw reads were generated, with 14,982,993 reads retained following the adapters and primers removal. After removal of reads found in positive and negative control samples, o 14,085,938 reads remained for further analysis. 3,263,040 unique reads were retained after dereplication and removal of singletons, and were used to identify 63 OTUs. 56 of these remained after running the LULU scripts. Although total number of sequencing reads retained in D2 was larger than that in ITS2, 11 samples (w2a109, w2a120, w2a220, w2b214, w2b314, w3b220, w4a309, w4a312, w4b112, w4b312 and w4b120) contained fewer than 1000 reads, thus were discarded from further analysis (Fig. S4 D2 total reads). Details of total number of reads and OTUs per sample in different extraction methods for both ITS2 and D2 are shown in Fig. S3 and Table S4.

Supplementary Tables

Table S1(A) Primers sequences used in this study and (B) Sequences of MiSeq sequencing adapters used in this study.

(A)

Targeted Region	**Primer name (Sequences 5' → 3')**	References
ITS2	fITS7_F (GTGAGTCATCGAATCTTTG)	Ihrmark et al., 2012
	ITS4_R (TCCTCCGCTTATTGATATGC)	White et al., 1990
D2	U1_F (GTGAAATTGTTGAAAGGGAA)	Putignani et al., 2008
	NL4_R (GGTCCGTGTTTCAAGACGG)	O'donnell, 1993

(B)

MiSeq sequencing adapters (5'-3')	
Fwd	AATGATACGGCGACCACCGAGATCTACACACACTGACGACATGGTTCTACA
	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA
Rev	CAAGCAGAAGACGGCATACGAGATTACGGTAGCAGAGACTTGGTCT

Table S2. The details of samples and summary of qPCR results in terms of cycles of threshold (C_t values) and number of copies in log of ITS2 and D2 markers.

No.	Sample name	Vineyard	Sampling date	Extraction Method	DNA concentration (ng/ul)	Yield (ng)	ITS2 C_t values	ITS2 log no. of copies	D2 C_t values	D2 log no. of copies
1	w2a109#	2	09-Oct	F	< 0.02*	1.10^	39.38	0.11	28.02	2.73
2	w2a209	2	09-Oct	C	1.34	46.90	32.65	2	21.26	4.8
3	w2a309	2	09-Oct	G	0.51	53.34	27.7	2.47	22.98	4.11
4	w2b109	2	09-Oct	F	0.28	15.18	37.27	0.83	23.66	4.17
5	w2b209	2	09-Oct	C	0.7	73.61	38.37	0.31	18.07	5.63
6	w2b309	2	09-Oct	G	0.76	79.59	38.21	0.35	16.22	6.12
7	w3a109	3	09-Oct	F	0.21	11.72	23.63	4.4	22.2	4.55
8	w3a209	3	09-Oct	C	8.72	305.2	19.81	5.59	16.04	6.16
9	w3a309	3	09-Oct	G	1.3	136.5	16.58	5.49	17.57	5.59
10	w3b109	3	09-Oct	F	0.45	24.48	21.46	5.12	19.2	5.15
11	w3b209	3	09-Oct	C	2.97	311.85	18.94	5.61	15.19	6.25
12	w3b309	3	09-Oct	G	5.3	556.5	20.05	5.31	14.5	6.56
13	w4a109	4	09-Oct	F	0.12	6.60	31.77	2.18	24.94	3.84
14	w4a209	4	09-Oct	C	0.41	14.21	36.43	0.94	18.81	5.44
15	w4a309#	4	09-Oct	G	0.3	31.29	23.55	3.6	20.9	4.89
16	w4b109	4	09-Oct	F	0.12	6.60	35.96	0.56	25.1	3.8
17	w4b209	4	09-Oct	C	0.26	27.30	37.13	0.65	19.43	5.28
18	w4b309	4	09-Oct	G	0.32	33.71	33.96	1.09	23.73	4.15
19	w5a109	5	09-Oct	F	< 0.02*	1.10^	38.38	0.38	27.53	2.86
20	w5a209	5	09-Oct	C	0.59	20.69	30.59	2.58	18.85	5.24
21	w5a309	5	09-Oct	G	0.18	18.59	29.46	1.99	21.84	4.65
22	w5b109	5	09-Oct	F	< 0.02*	1.10^	35.53	0.67	26.27	3.21
23	w5b209	5	09-Oct	C	0.21	21.63	25.15	3.92	20.1	5.11
24	w5b309	5	09-Oct	G	0.31	32.34	29.4	2.32	20.12	5.1
25	w2a112	2	12-Oct	F	0.22	12.21	34.54	1.43	25.19	3.77
26	w2a212	2	12-Oct	C	1.19	41.65	34.04	1.61	21.78	4.66
27	w2a312	2	12-Oct	G	0.21	21.53	33.83	0.81	25.16	3.78
28	w2b112	2	12-Oct	F	0.29	15.90	38.44	-0.11	20.51	4.79
29	w2b212	2	12-Oct	C	0.87	91.67	33.91	1.11	16.77	5.81
30	w2b312	2	12-Oct	G	1.14	119.7	33.63	1.18	18.36	5.56
31	w3a112	3	12-Oct	F	0.41	22.55	22.29	4.77	18.54	5.33
32	w3a212	3	12-Oct	C	2.8	98.00	24.4	4.31	17.95	5.67
33	w3a312	3	12-Oct	G	3.54	371.7	17.49	5.24	16.69	5.99
34	w3b112	3	12-Oct	F	0.29	15.90	21.83	4.35	21.69	4.69
35	w3b212	3	12-Oct	C	0.96	101.01	18.75	5.18	18.23	5.59

36	w3b312	3	12-Oct	G	5.96	625.8	16.65	5.74	16.82	5.96
37	w4a112	4	12-Oct	F	0.52	28.66	24.33	4.21	20.96	4.88
38	w4a212	4	12-Oct	C	5.26	184.1	22.84	4.74	16	6.02
39	w4a312 #	4	12-Oct	G	4.45	467.3	15.79	5.7	16.07	6.16
40	w4b112 §#	4	12-Oct	F	0.14	7.70	39.57	0.2	23.55	4.2
41	w4b212	4	12-Oct	C	0.41	43.16	37.3	0.82	18.97	5.4
42	w4b312 #	4	12-Oct	G	0.6	63.21	38.47	0.28	23.33	4.01
43	w5a112	5	12-Oct	F	< 0.02*	1.10^	29.43	2.82	23.12	4.07
44	w5a212	5	12-Oct	C	0.32	11.24	25.85	3.9	18.52	5.51
45	w5a312	5	12-Oct	G	0.27	27.83	22.2	3.96	22.34	4.52
46	w5b112	5	12-Oct	F	0.36	20.02	26.26	3.81	22.8	4.4
47	w5b212	5	12-Oct	C	0.86	90.72	23.88	4.26	17.3	5.83
48	w5b312	5	12-Oct	G	1.94	203.7	19.76	4.91	18.43	5.54
49	w2a114	2	14-Oct	F	< 0.02*	1.10^	38.73	0.29	30.87	2.29
50	w2a214	2	14-Oct	C	0.88	30.66	33.29	1.82	22.04	4.6
51	w2a314	2	14-Oct	G	0.23	24.36	36.07	0.2	24.55	3.94
52	w2b114	2	14-Oct	F	0.3	16.34	37.29	0.2	23.98	4.09
53	w2b214 #	2	14-Oct	C	0.93	98.07	33.13	1.32	17.57	5.76
54	w2b314 #	2	14-Oct	G	1.43	150.2	39.07	0.12	17.25	5.68
55	w3a114	3	14-Oct	F	0.19	10.45	26	3.76	22.42	4.5
56	w3a214	3	14-Oct	C	2.94	102.9	23.45	4.57	18.43	5.54
57	w3a314	3	14-Oct	G	2.69	282.5	16.6	5.48	16	6.17
58	w3b114	3	14-Oct	F	0.3	16.28	22.61	4.14	22.16	4.56
59	w3b214	3	14-Oct	C	0.69	72.66	18.59	5.22	17.42	5.64
60	w3b314	3	14-Oct	G	1.58	165.9	26.02	3.68	20.5	4.99
61	w4a114	4	14-Oct	F	0.47	25.80	24.37	4.2	20.53	4.99
62	w4a214	4	14-Oct	C	3.92	137.2	25.32	4.05	19.72	5.2
63	w4a314	4	14-Oct	G	4.04	424.2	16.38	5.54	15.77	6.23
64	w4b114	4	14-Oct	F	0.18	9.90	39.11	0.32	22.36	4.51
65	w4b214	4	14-Oct	C	0.48	49.98	33.69	1.8	19.2	5.34
66	w4b314	4	14-Oct	G	0.51	53.66	32.97	1.36	18.68	5.47
67	w5a114	5	14-Oct	F	0.21	11.33	25.75	3.82	20.94	4.88
68	w5a214	5	14-Oct	C	3.61	126.35	24.4	4.31	16.92	5.77
69	w5a314	5	14-Oct	G	1.19	124.95	15.69	5.73	15.75	6.09
70	w5b114	5	14-Oct	F	0.5	27.50	23.46	4.58	22.2	4.57
71	w5b214	5	14-Oct	C	2.03	213.2	19.83	5.56	17.1	5.9
72	w5b314	5	14-Oct	G	3.7	388.5	18.55	5.91	16.6	6.03

73	w2a120#	2	20-Oct	F	0.33	17.88	36.71	0.98	24.38	3.98
74	w2a220#	2	20-Oct	C	3.48	121.8	32.97	1.91	19.74	5.2
75	w2a320	2	20-Oct	G	3.48	365.4	27.8	2.45	19.99	5.13
76	w2b120	2	20-Oct	F	0.23	12.49	36.41	1.06	23.3	4.26
77	w2b220	2	20-Oct	C	0.65	68.15	35.36	1.34	19.3	5.31
78	w2b320	2	20-Oct	G	0.85	89.15	34.07	1.06	19.31	5.31
79	w3a120	3	20-Oct	F	< 0.02*	1.10^	27.94	3.36	25.65	3.38
80	w3a220	3	20-Oct	C	0.28	9.66	27.71	3.38	21.82	4.65
81	w3a320	3	20-Oct	G	0.13	13.23	18.58	4.95	21.08	4.85
82	w3b120	3	20-Oct	F	< 0.02*	1.10^	23.75	4.5	23.8	4.15
83	w3b220#	3	20-Oct	C	0.23	23.63	22.96	4.71	21.1	4.84
84	w3b320	3	20-Oct	G	0.36	37.70	20.58	4.69	21.31	4.79
85	w4a120	4	20-Oct	F	0.22	12.10	26.08	3.86	24.14	3.79
86	w4a220	4	20-Oct	C	0.3	10.33	27.84	3.35	22.53	4.47
87	w4a320	4	20-Oct	G	0.21	21.63	21.66	4.11	21.58	4.72
88	w4b120#	4	20-Oct	F	0.11	6.05	36.36	0.45	24.78	3.62
89	w4b220	4	20-Oct	C	0.18	18.90	37.83	0.06	20.51	4.79
90	w4b320	4	20-Oct	G	0.53	55.23	33	1.35	20.06	5.11
91	w5a120	5	20-Oct	F	< 0.02*	1.10^	24.27	4.35	23.5	4.22
92	w5a220	5	20-Oct	C	0.76	26.43	29.89	2.77	21.53	4.73
93	w5a320	5	20-Oct	G	0.75	78.23	17.22	5.31	17.74	5.72
94	w5b120	5	20-Oct	F	0.28	15.29	22.38	4.2	22.72	4.42
95	w5b220	5	20-Oct	C	0.2	21.00	22.1	4.28	20.8	4.71
96	w5b320	5	20-Oct	G	1.06	111.3	19.35	5.02	19.19	5.15

*Qubit value "< 0.02 ng/ul" means the concentrations were below detection limit.
^ Yield calculated based on concentration: 0.02 ng/ul.
§ Sample with fewer than 1000 total reads count and was removed from further analysis of ITS2.
# Samples with fewer than 1000 total reads count and was removed from further analysis of D2.

Table S3 A and B. The pairwise comparisons of (A) each extraction method using in terms of (i) DNA concentration (Qubit); (ii) cycles of threshold (C_t values) and (iii) log number of copies of D2 marker; (iv) cycles of threshold (C_t values) and (v) log number of copies of ITS2 marker. Also, pairwise comparisons (B) of each individual vineyard on the extraction methods (p-values shown) with the same parameters.

(A) Extraction methods

(i) Qubit	X^2= 35.198	df = 2	p-value = 2.27×10^{-8}
	Z value	p unadjusted	p adjusted
C vs F	5.138	2.78×10^{-7}	4.17×10^{-7}
C vs G	0	1.000	1.000
F vs G	-5.138	2.78×10^{-7}	8.3317×10^{-7}

(ii) D2 C_t values	X^2= 39.242	df = 2	p-value = 3.01×10^{-9}
	Z value	p unadjusted	p adjusted
C vs F	-5.674	1.39×10^{-8}	4.18×10^{-8}
C vs G	-0.538	0.590	0.590
F vs G	5.136	2.81×10^{-7}	4.22×10^{-7}

(iii) D2 no. of copies	X^2= 39.359	df = 2	p-value = 2.84×10^{-9}
	Z value	p unadjusted	p adjusted
C vs F	5.690	1.27×10^{-8}	3.81×10^{-8}
C vs G	0.556	0.578	0.578
F vs G	-5.134	2.84×10^{-7}	4.26×10^{-7}

(iv) ITS2 C_t values	X^2= 8.8959	df = 2	p-value = 0.0117
	Z value	p unadjusted	p adjusted
C vs F	-1.232	0.218	0.218
C vs G	1.737	0.082	0.124
F vs G	2.968	0.003	0.009

(v) ITS2 no. of copies	X^2= 4.7591	df = 2	p-value = 0.09259
	Z value	p unadjusted	p adjusted
C vs F	1.398	0.162	0.243
C vs G	-0.752	0.452	0.452
F vs G	-2.149	0.032	0.095

(B) Individual vineyards

(i) Qubit		Vineyard 2	Vineyard 3	Vineyard 4	Vineyard 5
	C vs F	0.001	0.010	0.040	0.013
	C vs G	0.400	0.798	0.505	0.878
	F vs G	0.031	0.018	0.040	0.010

(ii) D2 C_t values		Vineyard 2	Vineyard 3	Vineyard 4	Vineyard 5
	C vs F	0.001	0.003	0.001	0.000
	C vs G	0.645	0.645	0.574	0.959
	F vs G	0.028	0.005	0.021	0.001

(iii) D2 no. of copies		Vineyard 2	Vineyard 3	Vineyard 4	Vineyard 5
	C vs F	0.001	0.003	0.005	0.001
	C vs G	0.574	0.721	0.721	0.959
	F vs G	0.050	0.005	0.031	0.001

(iv) ITS2 C_t values		Vineyard 2	Vineyard 3	Vineyard 4	Vineyard 5
	C vs F	0.005	0.279	0.959	0.442
	C vs G	0.721	0.083	0.195	0.065
	F vs G	0.050	0.007	0.161	0.028

(v) ITS2 no. of copies		Vineyard 2	Vineyard 3	Vineyard 4	Vineyard 5
	C vs F	0.007	0.105	0.798	0.442
	C vs G	0.279	0.442	0.328	0.382

F vs G	0.293	0.021	0.442	0.083

Table S3 C and D. The p-values of pairwise comparisons of α-diversity indices: Observed OTUs, Chao1, Simpson and Shannon in ITS2 and D2 markers. The differences (C) between the DNA extraction methods were analysed using paired Wilcoxon signed-rank test (Wilcoxon, 1945) without adjusting for p-values, and (D) for pairwise extraction method and vineyard comparisons the familywise error rate was adjusted with Bonferroni correction (familywise error rate = 0.05).

C) Between extraction method

ITS2

	Observed OTUs	Chao1	Shannon	Simpson
C vs F	0.128	0.1529	0.005	0.009
C vs G	0.0001	0.0004	0.002	0.005
F vs G	0.003	0.0103	0.649	0.374

D2

	Observed OTUs	Chao1	Shannon	Simpson
C vs F	0.021	0.015	0.005	0.003
C vs G	0.026	0.066	0.420	0.346
F vs G	0.983	0.852	0.026	0.022

D) Pairwise extraction method and vineyard

ITS2

		Observed OTUs	Chao1	Shannon	Simpson
Vineyard 2	C vs F	0.115	0.272	0.109	0.313
	C vs G	0.014	0.008	0.008	0.008
	F vs G	0.023	0.035	0.742	0.383
Vineyard 3	C vs F	1	1	1	1
	C vs G	1	0.890	0.586	0.586
	F vs G	1	1	1	1
Vineyard 4	C vs F	0.279	0.341	0.093	0.093
	C vs G	0.074	0.106	0.800	0.933
	F vs G	0.752	0.833	0.208	0.142

Vineyard 5	C vs F	0.778	0.615	0.039	0.039
	C vs G	0.059	0.052	0.109	0.109
	F vs G	0.057	0.148	0.383	0.078
D2					
Vineyard 2		**Observed OTUs**	**Chao1**	**Shannon**	**Simpson**
	C vs F	0.577	0.581	0.855	0.855
	C vs G	0.170	0.279	0.438	0.438
	F vs G	0.583	0.583	0.813	0.813
Vineyard 3	C vs F	0.371	0.371	0.016	0.016
	C vs G	0.089	0.089	0.469	0.469
	F vs G	1	1	0.195	0.250
Vineyard 4	C vs F	1	0.773	0.156	0.156
	C vs G	0.773	0.773	0.100	0.1003
	F vs G	1	1	0.375	0.375
Vineyard 5	C vs F	0.058	0.058	0.195	0.109
	C vs G	0.281	0.430	0.742	0.844
	F vs G	0.598	0.674	0.078	0.055

Table S4(A) The OTUs abundance of ITS2 of each sample.

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w2a109	w2a112	w2a114	w2a120	w2a209	w2a212	w2a214	w2a220	w2a309	w2a312	w2a314	w2a320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	21158	57301	6268	8846	53894	65113	55344	54936	119873	28051	59982	43031
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	18	0	0	71	20	0	174	111
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0	18
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	27	2	0	55	132	41	149	260	512	537	653
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	26
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	21	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	139	63	167	3	52	36	53	88	614	2549	256
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	14	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	0	1	0	0	2	1	49	1	3	0	11	15
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	2	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	17	0	0	50
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	1	0	0	0	0	0	0	0	0	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	1	0	0	0	5	0	0	6
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	114	0	0	0	0	0	2	57	214	98	86
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	106	4	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	21	0	16	0	0	22
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	1	1	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	88	0	0	0	69	0	14	38	323	35	11
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	1029	2101	399	37215	3862	22256	127376	19229	7230	2973	6611	9928
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	1	0	0	1	23	9	43	1	0	0	19
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	0	0	0	146	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	16	0	0	65	0	0	0	30
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	4	11	1	14	18	62	74	5
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	1	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	1	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	1251	0	0	127	264	338	766	1379	3783	15098	4321
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	1	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	20
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	341	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	1317	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	0	43	0	0	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	0	7	495	0	0	0	0	0	0	1	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	178	0	62	347	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	1183	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	3	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	11	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	171	0	0	21	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	10	0	3	0	0	28	2	5	44	32	41
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	1	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	3	0	0	0	0	0	1	0	111	3	19

OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	340	0	1195	114	391	11	222	741	477	1439	1511
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	296	0	1	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	375	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	327	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	85	9
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	250	2	0	0	0	19	0	34	0	133	44
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	2	0	0	0	0	0	0	0	0	0	5
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	277	1113	0	1487	27	12	16	430	274	216	638	592
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	41	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	77	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	1	0	0	0	0	0	24	0	3	68
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	1	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	85	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	45	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w2b109	w2b112	w2b114	w2b120	w2b209	w2b212	w2b214	w2b220	w2b309	w2b312	w2b314	w2b320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	5	4014	42	32	1	42	483	0	705	1	4731	126
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	1	0	0	0	0	0	0	0	3	1548	1193	120
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	30
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	2	1410	0	0	800	6	264	0	5190	1395	3391	243
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	25	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	2	0	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	415	1	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	115	16	283	163	2134	204	104	410	889	90	234	514
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	123	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	279	27	57	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	855	1	1228	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	1	0	0	1	2	0	0	11
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	0	507	283	2150	128
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	4931	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	4261	0	0	0	0	0	0	391	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	560	0	0	0	1276	257	710	2
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	1895	7401	9129	98644	24061	61569	30129	186624	21732	9302	80418	31726
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	29	0	971	17	1230	17

OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	741	490	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	211	0	868	150	1098	185
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	707	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	447	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	454	35	68	5	214	1	55	18	608	159	206	4
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	61	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	1750	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	46	1	1739	0	5250	104	805	1779	42327	8981	44034	3123
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	1	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	1	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	103	531	0	54	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	331	194	571	27	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	908	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	848	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	0	78	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	6	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	626	0	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	845	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	4	4	0	4	413	41	12	0	420	107	896	316
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	579	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	697	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	475	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	434	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	516	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	51	2	69	20	119	19	107	1
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	1	0	0	0	823	29	900	0	25051	5476	16083	756
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	319	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	445	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	236	1	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	117
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	38	1154	865	0	1158	0	420	0	3031	31950	2110	594
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	160	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	13	0	3	11	1	19	9	1
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	214	733	737	321	1878	16	885	0	11441	3100	11839	776
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	35	62	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	89	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	1124	0	0	0	0	0	0	8	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	1	0	0	0	2	8	0	0	0	3	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w3a109	w3a112	w3a114	w3a120	w3a209	w3a212	w3a214	w3a220	w3a309	w3a312	w3a314	w3a320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	208245	196340	125230	83971	105800	109071	115134	198020	136698	220105	216318	88740
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	1	0	0	0	0	2	3	0	0	0	0	0
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	1	3	1	1	0	0
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	2	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	2	0	0	0	0	0	0	0	0	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	1	1	0	0	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	1	0	0	0	3
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	2	1	1	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	16	0	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	1	0	1	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	0	0	0	0	1	19	0	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	0	6	0	0	0	0
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	0	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	5	2	0	0	1	2	4	24	3	1	2	1
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	3	0	0	0	0	0	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	0	0	0	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	0	0	0	0	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	1	0	0	0	0	0	0	0	0	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	2	0	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	1	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	26	0	0	0	2	5	2	4	0	9	10	0
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0

OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	3	1	0	0	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	5	1	0	0	0	0	0	0	0	0	0	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	4	0	1	1	0	1	0	0	1	0	0	12
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w3b109	w3b112	w3b114	w3b120	w3b209	w3b212	w3b214	w3b220	w3b309	w3b312	w3b314	w3b320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	99058	131669	94482	188340	179306	101236	231651	89138	187194	70045	71368	122983
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	15	1	0	0	0	0	0	0	0	0
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	1	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	1	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	0	0	3	0	2	0
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	1	0	1	0	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	1	2	1	1	0	0	1	1	5
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	1	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	5	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	48	0	0	0	1	0	187	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	0	0	0	0	0	0

OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	0	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	1	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	1	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	2	3	1	4	5	1	2	8	6	4	15
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	1	0	0	0	0	0	0	0	0	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	0	0	0	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	0	0	0	0	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	1	9	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	8	3	5	11	9	7	0	10	12	11	6
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	1	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	2	0	0	0	1	0	0	0	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	1	0	0	0	1	0	0	0	2	0	0	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	1	2	1	0	0	1	0	3	1	1	0
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	2	4	0	2	3	2	8	7	7	3	5	8
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w4a109	w4a112	w4a114	w4a120	w4a209	w4a212	w4a214	w4a220	w4a309	w4a312	w4a314	w4a320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	47722	108123	70854	52869	130464	128958	128990	88747	40732	99973	87017	134571
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	1	0	0	7	4	0	0	17	5	0	0
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	26	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	12	0	0	0	1	0	1	0	5	0	1	2
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	3	0	0	0	8	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	1	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	1	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	2	0	0	0	0	1	1	0	0	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	1	1	0	4	0	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	0	1	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	1	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	48	0	0	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	1	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	10	0	0	0	0	0	0	0	0	0	0	0
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	24	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	20	3	0	5	68	17	9	27	204	21	12	27
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	1
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	31	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	13	0	0	0	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	6	0	1	0	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	1	0	0	1
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	77	0	0	0	1	0	1	0	0	0	0	0
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0

OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	1	0	0	0	2	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	1
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	1	0	2	0	1	1	0	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	1	0	2	5	1	0	0	0	2	0	1	0
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	37	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w4b109	w4b114	w4b120	w4b209	w4b212	w4b214	w4b220	w4b309	w4b312	w4b314	w4b320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	16139	112	91	529	0	0	1	55	526	49	172
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	1216	2559	0	0	401	0	350	302	4194	1109	694
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	40	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	24
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	0	0	2370	2085	2167	1211	770	2423	2414	1623
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	24	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	17	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	19	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	17	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	1	0	0	0	3	5	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	38	332	634	691	2608	1324	2660	124	40	19	922
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	516	0	0	0	21	62	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	1803	3272	1501	100	330	120	1479
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	67	362	279	125
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	13	0	0	3	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	27	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	5412	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	717	0	519	0	166	119	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	388	1892	804	4431	6834	28078	71	240	133	8923
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	351	15	18	215	202	113

OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	287	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	2676	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	1925	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	28	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	4827	182	10	27979	47397	54074	4940	10326	67493	31340	21448
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	76	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	1104	0	0	0	0	1137	0	75	234	115	97
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	1271	0	0	0	0	0	0	0	0	1	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	1397	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	12	0	155	106	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	6	0	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	62	429	348	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	1345	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	914	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	803	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	784	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	318	165	209	134
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	166	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	662	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	2	17	2	0	0	1
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	579	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	145	0	227
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	138	0	0	0	0	0	0
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	2618	0	1890	12	3655	0	318	346	258	201
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	198	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	190	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	1528	319	533	1411	2425	1103	538
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	213	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	126	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	101	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	429	2513	0	1007	1656	320	422	294	1561	1387	605
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	44	47	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	1	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	79	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	14	11	0	0	3	31	0	0	2	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	56	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	48	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	46	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	38	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	33	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w5a109	w5a112	w5a114	w5a120	w5a209	w5a212	w5a214	w5a220	w5a309	w5a312	w5a314	w5a320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	15	27503	245647	132545	191	168314	182449	163635	75	59112	157291	93873
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	2	129	8	0	507	1	7	0	18290	197	49	6
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	23	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	49	0	0	0	0	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	0	0	0	12	0	0	0	0	0	0	0
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	38	0	1	1	8222	0	0	6	230	3	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	3	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	1	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	0	2	0	0	21	1	0	0	0	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	39	0	0	0	449	4	1	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	40	0	0	0	98	4	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	1	0	0	0	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	148	0	0	5	1265	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	2	0	0	0	0	0	93	0	1	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	0	0	65	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	0	0	0	0	0	0	2	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	0	0	129	0	0	0
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	262	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	17	4	1	0	24	6	1	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	1529	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	3491	13	16	2	10990	7	10	50	24624	190	22	6
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	3	0	0	0	3	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	42	0	0	0	1588	10	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	0	0	1	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	313	1	0	1	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	37	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	1	0	0	0	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	8528	6421	1417	3	70240	417	220	40	62026	1176	44	47
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	461	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	4	0	0	0

OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	42	0	0	0	0	6	0	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	4303	1	0	0	348	0	4	0	3170	44	8	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	129	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	219	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	1	0	0	0	24	0	0	0	7	1	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	1005	4	11	1	1291	15	2	0	4288	122	9	3
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	103	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	1	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	63	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	3509	5	5	2	20777	15	5	50	11784	99	4	2
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w5b109	w5b112	w5b114	w5b120	w5b209	w5b212	w5b214	w5b220	w5b309	w5b312	w5b314	w5b320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	1446	116186	152211	235431	92	145390	256066	230417	5495	72042	206915	111721
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	261	41	6	15	32	0	2	1	642	15	9	1
OTU_100	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Lodderomyces	Lodderomyces	0	0	0	0	8	0	0	0	0	0	0	0
OTU_101	Fungi	NA	NA	NA	NA	NA	Fungi OTU_101	0	0	0	0	0	0	0	0	0	0	0	0
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_104	Fungi	Basidiomycota	Tremellomycetes	Cystofilobasidiales	NA	NA	Cystofilobasidiales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_105	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Dioszegia	Dioszegia	0	0	0	0	0	0	0	0	5	0	0	0
OTU_106	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Parastagonospora	Parastagonospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_107	Fungi	Ascomycota	Dothideomycetes	Pleosporales	NA	NA	Pleosporales	0	0	0	0	0	0	0	0	3	1	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	4	0	0	5	0	0	0	19	2	0	0
OTU_110	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_111	Fungi	NA	NA	NA	NA	NA	Fungi OTU_111	0	0	0	0	0	0	0	0	0	0	0	0
OTU_112	Fungi	NA	NA	NA	NA	NA	Fungi OTU_112	0	0	0	0	0	0	0	0	0	0	0	0
OTU_113	Fungi	NA	NA	NA	NA	NA	Fungi OTU_113	0	0	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_115	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_116	Fungi	Basidiomycota	Agaricomycetes	Hymenochaetales	Schizoporaceae	Hyphodontia	Hyphodontia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	5	0	0	0
OTU_120	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Phaffomycetaceae	Wickerhamomyces	Wickerhamomyces	0	0	0	0	0	0	0	0	5	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	302	0	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	1	1	0	11	0	0	0	0	0	0	0
OTU_126	Fungi	Ascomycota	Sordariomycetes	Diaporthales	Valsaceae	Cytospora	Cytospora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_129	Fungi	NA	NA	NA	NA	NA	Fungi OTU_129	0	0	0	0	0	0	0	0	10	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	1	0	0	0	17	1	0	0	4	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	1	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	4	0	0	0	0	0	0	53	1	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	13	19	0	0	0	71	2	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycetes	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	6	0	0	41	0	0	0	507	2	1	1
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_163	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	NA	Peniophoraceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	30	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	0	0	6	1	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	0	0	0	2	0	0	69	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	0	0	0	0	0	0	0	8	1	0	0

OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	1627	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	4	0	0	6	0	1	0	152	1	1	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	1	0	0	0	0	35	1	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	1753	0	0	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	19790	97	13	55	184623	11	5	11	16120	115	31	5
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	1704	0	0	0	0	0	0	0	5	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	1	0	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	1330	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	1	0	0	0	0	268	1	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	2379	0	0	0	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	1121	3	0	0	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_41	Fungi	NA	NA	NA	NA	NA	Fungi OTU_41	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_44	Fungi	NA	NA	NA	NA	NA	Fungi OTU_44	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	7	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	480	0	0	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	1	0	0	0	0	0	0	0	0
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	9744	866	13	46	12610	15	7	71	6200	169	23	127
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	1	0	0	0	0	0	0	0	0	0	0
OTU_51	Fungi	Ascomycota	Dothideomycetes	Botryosphaeriales	Botryosphaeriaceae	Diplodia	Diplodia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_55	Fungi	Basidiomycota	Agaricomycetes	Russulales	Peniophoraceae	Peniophora	Peniophora	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_58	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	3	16	0	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_62	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0
OTU_67	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Debaryomycetaceae	Debaryomyces	Debaryomyces	0	0	0	0	0	0	0	0	0	0	0	0
OTU_68	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_69	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	1088	14	5	0	5	0	0	2	792	2	2	1
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycetes	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_74	Fungi	Ascomycota	Sordariomycetes	Xylariales	Xylariaceae	Lopadostoma	Lopadostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_75	Fungi	NA	NA	NA	NA	NA	Fungi OTU_75	0	0	0	0	0	0	0	0	0	0	0	0
OTU_77	Fungi	NA	NA	NA	NA	NA	Fungi OTU_77	0	0	0	0	0	0	0	0	8	0	0	0
OTU_78	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	NA	Phaeosphaeriaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	3	0	0	0	0	0	0	17	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	2346	34	7	5	67	0	5	5	8872	36	11	1
OTU_80	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Melanommataceae	NA	Melanommataceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_82	Fungi	NA	NA	NA	NA	NA	Fungi OTU_82	0	0	0	0	0	0	0	0	0	0	0	0
OTU_83	Fungi	NA	NA	NA	NA	NA	Fungi OTU_83	0	0	0	0	0	0	0	0	0	0	0	0
OTU_84	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Allophaeosphaeria	Allophaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_85	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Dissoconiaceae	Dissoconium	Dissoconium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_86	Fungi	Basidiomycota	Agaricostilbomycetes	NA	NA	NA	Agaricostilbomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	70	0	0	0	0	0	0	0	0	0	0	0
OTU_89	Fungi	NA	NA	NA	NA	NA	Fungi OTU_89	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	3549	3	0	0	73	1	1	1	415	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0
OTU_92	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Phaeosphaeriaceae	Phaeosphaeria	Phaeosphaeria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	1	0	0	0	0	0	0	0	0	0	0
OTU_94	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0
OTU_95	Fungi	NA	NA	NA	NA	NA	Fungi OTU_95	0	0	0	0	0	0	0	0	0	0	0	0
OTU_96	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Naganishia	Naganishia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_99	Fungi	NA	NA	NA	NA	NA	Fungi OTU_99	0	0	0	0	0	0	0	0	0	0	0	0

Table S4(B) The OTUs abundance of D2 of each sample.

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w2a112	w2a114	w2a209	w2a212	w2a214	w2a309	w2a312	w2a314	w2a320	w2b109	w2b112	w2b114
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	27038	6112	58004	128461	26958	34740	64015	53355	59760	94016	90142	115588
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	2	4	0	1	0	0	0	3	1	0	0	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	1	17	0	0	0	0	0	0	0	0	0	1
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	2	0	10	0	1	6	5	4	9	0	1	2
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	2	0	0	0	0	0	0	0	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	3	3	0	0	1	0	3	3	2	3	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	707	640	534	2	153	545	275	184	196	1	4	34
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	1	0	0	0	0	10	6	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccotheciaceae	Aureobasidium	Aureobasidium	1	0	0	7627	0	1	11	9	6	2	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	28	371	1	0	15	4	90	138	7	0	0	0
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	25	0	0	0	0	0	0	0	0	0	7
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	1	0	0	1	0	0	0	0	0	0	0	0

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w2b120	w2b209	w2b212	w2b220	w2b309	w2b312	w2b320	w3a109	w3a112	w3a114	w3a120	w3a209
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	50210	131725	110847	73661	150058	121815	66336	28575	110	228	68722	90977
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	1	0	0	0	0	1
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	1	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	13	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	2	5	0	0	4	7	2	0	0	0	0	0
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	2	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	27	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	0	0	0	0	0	1	1	0	0	0	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	2
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	6	5	49	4	117	123	151	129794	99175	154499	100625	120935
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	0	0	0	1
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccotheciaceae	Aureobasidium	Aureobasidium	0	0	0	0	0	0	3	0	0	1	1	1
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	14	1	0	0	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	4	0
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	225	1	0	0	1
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	0	0	0	0	0	0	4	0

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w3a212	w3a214	w3a220	w3a309	w3a312	w3a314	w3a320	w3b109	w3b112	w3b114	w3b120	w3b209
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	36006	49219	63600	22819	76330	118937	12432	236	385	4852	1847	29572
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	1	1	0	1	0	3	0	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	1
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	1	18	0	2
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	5	3	0	7	0	2	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	2	0	2	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	0	1	0	0	1	1	0	0	0	0	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	127199	180485	85373	139847	425870	149420	164055	105672	133756	168961	127003	134705
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	0	0	0	0

OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	2	0	0	0	0	0	0	1	3	1	0
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccothecіaceae	Aureobasidium	Aureobasidium	0	0	2	0	11	2	1	0	2	2	1	1
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	3	2	0	2	1	6	0	3	54	0	82	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	9	0	0
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	6	4	6	80	0	15	0	0	2	0	0	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	0	0	0	0	0	0	0	0	0	0	1
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	1	1	0	0	0	1	0	0

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w3b212	w3b214	w3b309	w3b312	w3b314	w3b320	w4a109	w4a112	w4a114	w4a120	w4a209	w4a212
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	8017	7339	2674	18177	17795	28364	110163	10714	146	45	107998	3353
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	1	2	1	2	0	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	1	0	1	0	0	0
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	3
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	2	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	1	0	1	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	6	5	2	5	4	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	144854	154054	120762	126888	126933	127239	7672	133457	141350	113325	21132	162698
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	1	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	1	0	1	0	0	0	0	0	0	0
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccothecіaceae	Aureobasidium	Aureobasidium	0	0	2	1	2	5	1	7	0	4	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	635	269	109	1235	924	0	4	0	0	0	1	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	1
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	1	0	0	1	0	0	0	0	0	0	0	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	0	0	0	0	0	8	0	1	0	0	0
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	0	0	0	3	0	6	0	0

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w4a214	w4a220	w4a314	w4a320	w4b109	w4b114	w4b209	w4b212	w4b214	w4b220	w4b309	w4b314
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	63482	62979	36587	9593	118282	150334	149902	100045	81824	128882	145312	157438
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	1	0	0	0	2	0	2	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	1	0	0	0	0	0	0	1	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	0	0	0	0	1	0	1	2	0	2	0	0
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	2	0	0	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	4	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	4	0	4	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	0	2	1	0	0	0	0	0	0	1	5	2
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	90040	98990	167615	104291	105	0	41	3	94998	13	41	72
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	0	0	0	2	1
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	0	0	1	0	0	0
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccothecіaceae	Aureobasidium	Aureobasidium	2	3	1	2	1	1	2	1	3	3	8	11
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	1	0	0	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	1	0	0	6	0	0	0
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	1	32	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	1	0	0	0	4	2	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	0	0	0	0	1	0	0	0

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w4b320	w5a109	w5a112	w5a114	w5a120	w5a209	w5a212	w5a214	w5a220	w5a309	w5a312	w5a314
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	142149	15778	4627	575	243	99834	55494	158140	22634	105130	45115	2966
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	9	9	0	0	0	0	0	2	0	7	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	8	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	3	0	0	0	0	0	0	4	0	4	0	0

OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	36	0	0	0	4	6	0	6	0	2	0
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	9	0	7	0	23	3
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	1	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	3	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	3	0	0	1	0	0	0	0	0	0	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	28	0	0	0	0	3	0	0	0	1	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	12	154	113980	143249	127969	76	57054	43	134494	16	44793	139304
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	1	0	0	0	0	0	0	0	5	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	18	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	14	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	0	0	0	0	0	0	3	0	0	0	11	1
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	1	0	0	0	0	0	0	1	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	0	1	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccotheciaceae	Aureobasidium	Aureobasidium	16	82	118	0	4	6	1	0	8	3	14	4
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	77	0	0	0	0	0	0	0	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	20	0	0	0	269	1	0	10	0	7	0
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	68	0	0	2	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	1	0	87	0	1	1	1	0	0	0	15	1

OTU#	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w5a320	w5b109	w5b112	w5b114	w5b120	w5b209	w5b212	w5b214	w5b220	w5b309	w5b312	w5b314	w5b320
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	23362	67540	14038	10895	11104	42358	19070	15029	22514	130983	99644	30392	6856
OTU_10	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	1	53	8	0	0	1	0	0	0	17	4	0	0
OTU_11	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	99	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	NA	Saccharomycodaceae	0	0	2	0	5	3	0	0	0	0	0	0	0
OTU_14	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	1	0	0	2
OTU_15	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	3	0	0	1	0	0	0	1	0	0	21	4	18
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	4	0	0	0	0	3	0	0	0	0
OTU_18	Fungi	Basidiomycota	Tremellomycetes	Filobasidiales	Filobasidiaceae	Filobasidium	Filobasidium	1	0	0	0	0	0	0	0	0	1	0	0	0
OTU_19	Fungi	Mucoromycota	NA	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	2	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	143671	34	584	191550	92	174	245898	150042	142103	84	108755	129699	160481
OTU_20	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	19	0	0	0	0	0	0	0	1	1	0	1
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Physalacriaceae	Cylindrobasidium	Cylindrobasidium	0	9	0	0	0	0	0	0	0	0	0	0	0
OTU_24	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	NA	Metschnikowiaceae	1	0	0	0	0	0	0	0	0	0	4	2	3
OTU_25	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Sordariomycetes	Ophiostomatales	Ophiostomataceae	Ophiostoma	Ophiostoma	0	0	10	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	NA	Trichosporonaceae	0	0	14	0	0	0	0	0	0	0	0	0	0
OTU_4	Fungi	Ascomycota	Dothideomycetes	Dothideales	Saccotheciaceae	Aureobasidium	Aureobasidium	8	72	0	0	1	3	0	0	1	25	14	4	4
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	524	0	0	1	0	313	0	0	0	0
OTU_6	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	1	1	2	0	1	7	0	0	0	1	1	1	5
OTU_7	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Pichiaceae	Pichia	Pichia	0	0	0	0	0	0	16	0	0	0	0	1	0
OTU_8	Fungi	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia	Malassezia	0	87	26	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	0	0	0	0	0	0	3	0	3

Table S5. Results of (A) PERMANOVA of the effect of vineyard and extraction methods, and their interactions on *β*-diversity; (B) multivariate dispersion (PERMDISP2) of different extraction methods (F, C & G); (C) pairwise PERMANOVA comparison of the effect of vineyard and extraction methods, as well as their interactions. All analyses were done for both D2 and ITS markers under different distance metrics.

(A) PERMANOVA

Marker	Distance metric	Vineyard			Extraction.Method			Vineyard x Extraction.Method		
		R^2	pseudo-F	p-value	R^2	pseudo-F	p-value	R^2	pseudo-F	p-value
D2	Bray-Curtis	0.239	9.002	0.001	0.071	3.983	0.002	0.044	0.830	0.722
	Jaccard	0.173	5.950	0.001	0.059	3.038	0.002	0.061	1.053	0.382
	Poisson	0.289	11.495	0.001	0.067	3.974	0.025	0.034	0.672	0.793
ITS2	Bray-Curtis	0.311	16.295	0.001	0.097	7.592	0.001	0.058	1.645	0.047
	Jaccard	0.218	9.541	0.001	0.076	5.017	0.001	0.075	1.637	0.001
	Poisson	0.186	10.093	0.001	0.185	15.053	0.001	0.121	3.287	0.001

(B) PERMDISP2

Marker	Distance metric	ANOVA		Tukey multiple comparisons of pairwise means adjusted p-value		
		F-value	p-value	F vs C	G vs C	F vs G
D2	Bray-Curtis	6.486	0.002	0.025	0.703	0.003
	Jaccard	3.288	0.042	0.755	0.178	0.039
	Poisson	3.378	0.039	0.040	0.854	0.138
ITS2	Bray-Curtis	4.938	0.009	0.034	0.938	0.014
	Jaccard	4.445	0.014	0.031	0.999	0.029

Poisson	1.864	0.161	0.136	0.561	0.636

(C) PAIRWISE-PERMANOVA

D2 marker

Extraction method	Distance metric	Vineyard	Extraction.Method	Vineyard x Extraction.Method					
		R^2	pseudo-F	p-value	R^2	pseudo-F	p-value	R^2	pseudo-F	p-value
F vs C	Bray-Curtis	0.203	4.579	0.001	0.050	3.412	0.007	0.024	0.533	0.932
	Jaccard	0.142	2.944	0.001	0.035	2.161	0.034	0.037	0.779	0.761
	Poisson	0.229	5.277	0.001	0.052	3.620	0.026	0.008	0.195	0.964
G vs C	Bray-Curtis	0.294	7.794	0.001	0.056	4.464	0.002	0.033	0.880	0.573
	Jaccard	0.218	5.191	0.001	0.053	3.752	0.001	0.044	1.037	0.412
	Poisson	0.456	14.956	0.001	0.025	2.468	0.107	0.022	0.721	0.592
F vs G	Bray-Curtis	0.276	7.097	0.001	0.056	4.294	0.002	0.045	1.157	0.297
	Jaccard	0.218	5.188	0.001	0.048	3.430	0.001	0.060	1.414	0.103
	Poisson	0.326	8.770	0.001	0.051	4.126	0.018	0.029	0.794	0.586

(C) PAIRWISE-PERMANOVA (cont'd)

ITS2 marker

Extraction method	Distance metric	Vineyard	Extraction.Method			Vineyard x Extraction.Method					
			R^2	pseudo-F	p-value	R^2	pseudo-F	p-value	R^2	pseudo-F	p-value
F vs C	Bray-Curtis		0.306	9.981	0.001	0.083	8.171	0.001	0.049	1.595	0.047
	Jaccard		0.213	5.905	0.001	0.065	5.398	0.001	0.059	1.629	0.010
	Poisson		0.243	7.719	0.001	0.121	11.596	0.001	0.060	1.905	0.061
G vs C	Bray-Curtis		0.421	15.431	0.001	0.033	3.657	0.008	0.036	1.314	0.196
	Jaccard		0.312	9.499	0.001	0.032	2.894	0.006	0.044	1.342	0.108
	Poisson		0.466	18.605	0.001	0.038	4.496	0.083	0.028	1.113	0.355
F vs G	Bray-Curtis		0.294	9.835	0.001	0.100	9.991	0.001	0.058	1.947	0.011
	Jaccard		0.206	5.782	0.001	0.076	6.394	0.001	0.067	1.888	0.002
	Poisson		0.100	3.486	0.023	0.231	24.233	0.001	0.145	5.063	0.002

Table S6. The differential abundance analysis of OTUs in (A) D2 and (B) ITS2 markers in different extraction method comparisons. All the effects with adjusted p-value less than 0.1 are shown.

A) D2

G vs C

	BaseMean	log2 FoldChange	lfcSE	stat	p-value	p-adj	Taxonomic annotation
OTU_1	262113	4.236	1.133	3.739	0.0002	0.002	Hanseniaspora
OTU_5	80.0	14.18	3.579	3.962	0.0002	0.002	Pichia

B) ITS2

F vs G

	BaseMean	log2 FoldChange	lfcSE	stat	p-value	p-adj	Taxonomic annotation
OTU_10	40.20	10.295	1.817	5.665	1.47×10^{-8}	3.97×10^{-7}	Botrytis
OTU_11	17.21	3.962	1.662	2.384	0.017	0.051	Cryptococcus
OTU_14	5.761	6.451	2.529	2.551	0.011	0.036	Vishniacozyma
OTU_20	1.488	6.910	2.641	2.617	0.009	0.034	Cryptococcus
OTU_22	3.422	8.483	2.779	3.052	0.002	0.012	Kazachstania
OTU_3	228	4.102	0.951	4.316	1.59×10^{-5}	1.76×10^{-4}	Dothideomycetes
OTU_38	4.711	-12.09	2.832	-4.270	1.96×10^{-5}	1.76×10^{-4}	Fungi OTU_38
OTU_5	789	4.486	1.406	3.191	0.001	0.010	Metschnikowia
OTU_6	26.96	4.162	1.989	2.092	0.036	0.098	Papiliotrema
OTU_7	43.73	4.842	1.684	2.874	0.004	0.018	Alternaria

F vs C

	BaseMean	log2 FoldChange	lfcSE	stat	p-value	p-adj	Taxonomic annotation
OTU_10	40.20	9.168	1.824	5.027	4.97×10^{-7}	4.62×10^{-5}	Botrytis
OTU_20	1.488	7.314	2.650	2.760	0.006	0.090	Cryptococcus

OTU_22	3.422	9.434	2.783	3.390	0.001	0.019	Kazachstania
OTU_3	228	3.189	0.951	3.353	0.001	0.019	Dothideomycetes
OTU_38	4.711	-9.014	2.830	-3.185	0.001	0.027	Fungi OTU_38
OTU_5	789	5.895	1.410	4.181	2.90×10^{-5}	0.001	Metschnikowia

Supplementary Figures

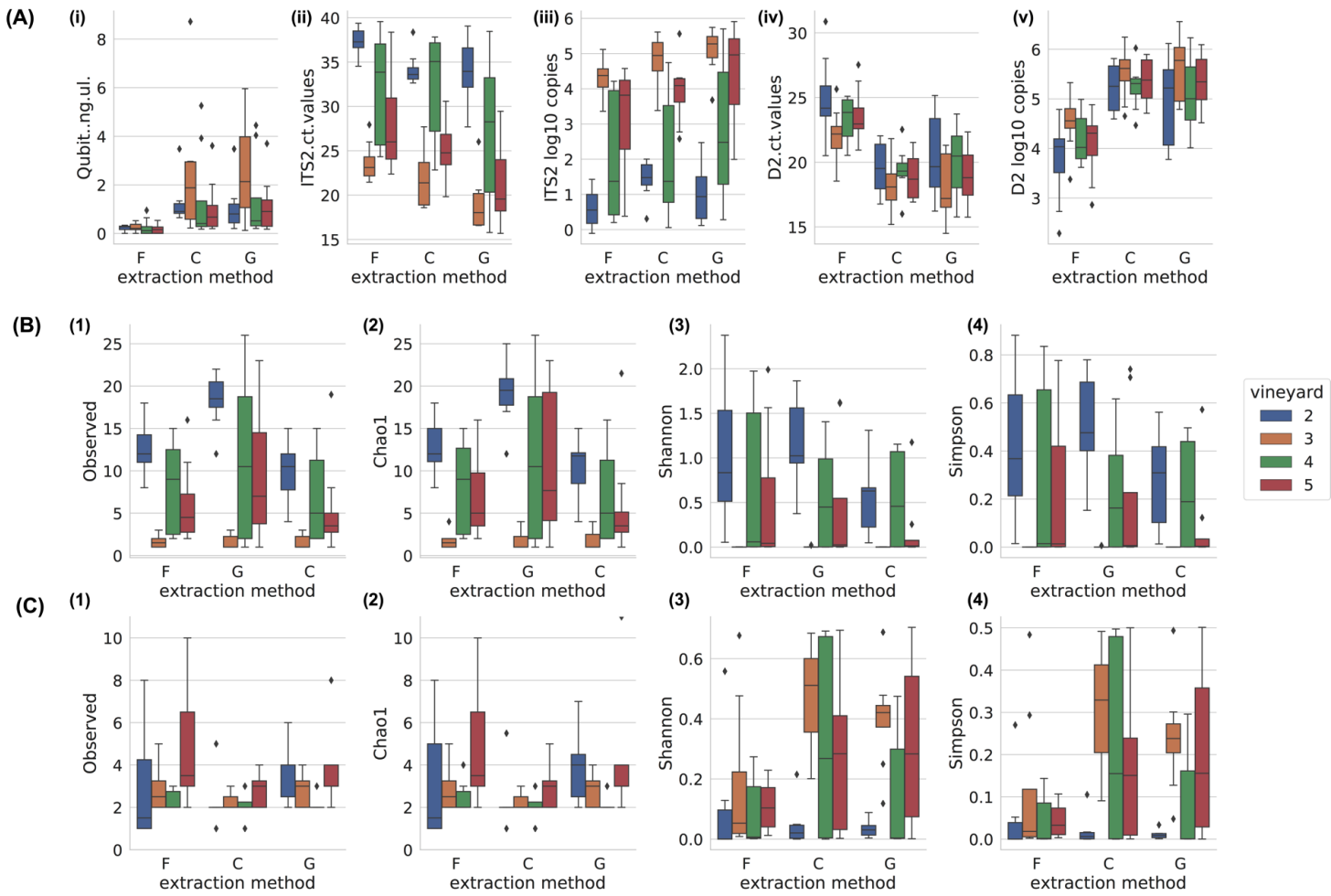


Fig. S1(A) The DNA concentration (Qubit values), the cycles of threshold (C_t values), log number of copies among extraction methods (F, C and G) and in different vineyards (2: blue; 3: organe; 4: green; 5: red) for ITS2 and D2 markers. (i) the DNA concentrations from Qubit HS kit; (ii) and (iii) C_t values and the log number of copies of ITS2 markers; (iv) and (v) C_t values and the log number of copies of D2 markers. S1B

and C) showed the species richness and alpha-diversity of fungal community of two markers. (1) Observed OTUs, (2) Chao1 (predicted OTUs), (3) Shannon index and (4) Simpson index of ITS2 (B) and D2 (C) markers per extraction method.

(A) **ITS2**

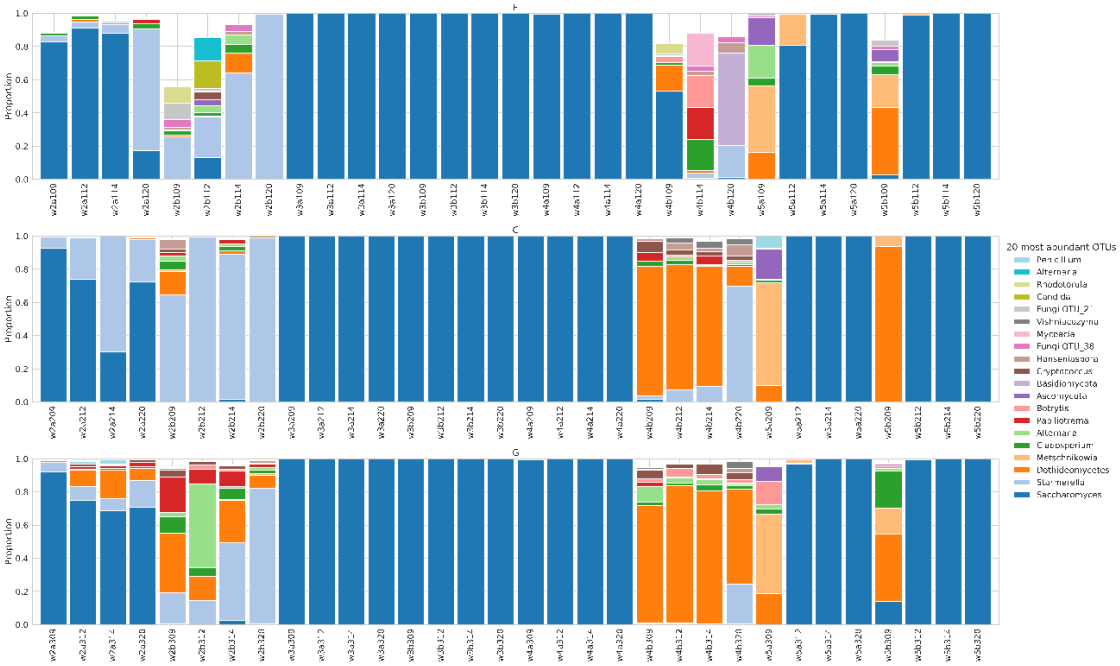


Fig. S2A Relative abundance of top 20 fungal OTUs from ITS2 markers of different extraction methods F (top), C (middle) and G (bottom) per sample. Sample w4b112 in method F (top) was removed owing to the total reads counts < 1000. *Saccharomyces* was most abundant in total number of reads. White proportion equal to the rest of OTUs.

(B) D2

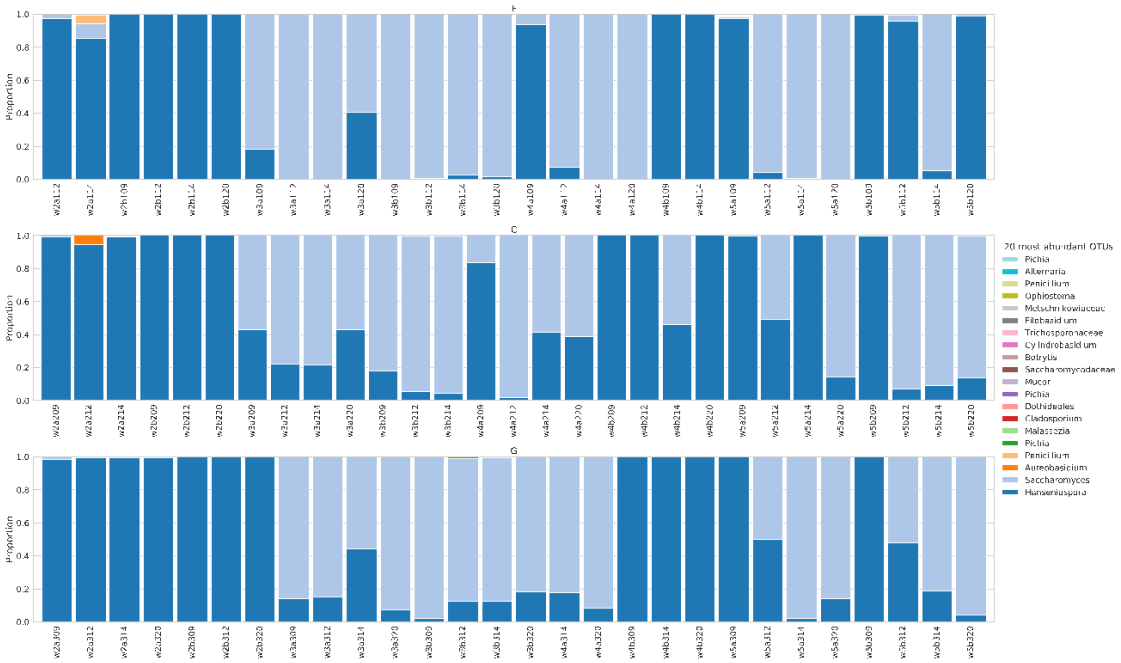


Fig. S2B Relative abundance of top 20 fungal community from D2 markers of different extraction methods F (n= 28, top), C (n= 29, middle) and G (n=28, bottom) per sample. Eleven samples were removed owing to the total reads counts < 1000. Hanseniaspora and Saccharomyces was dominated in total number of reads.

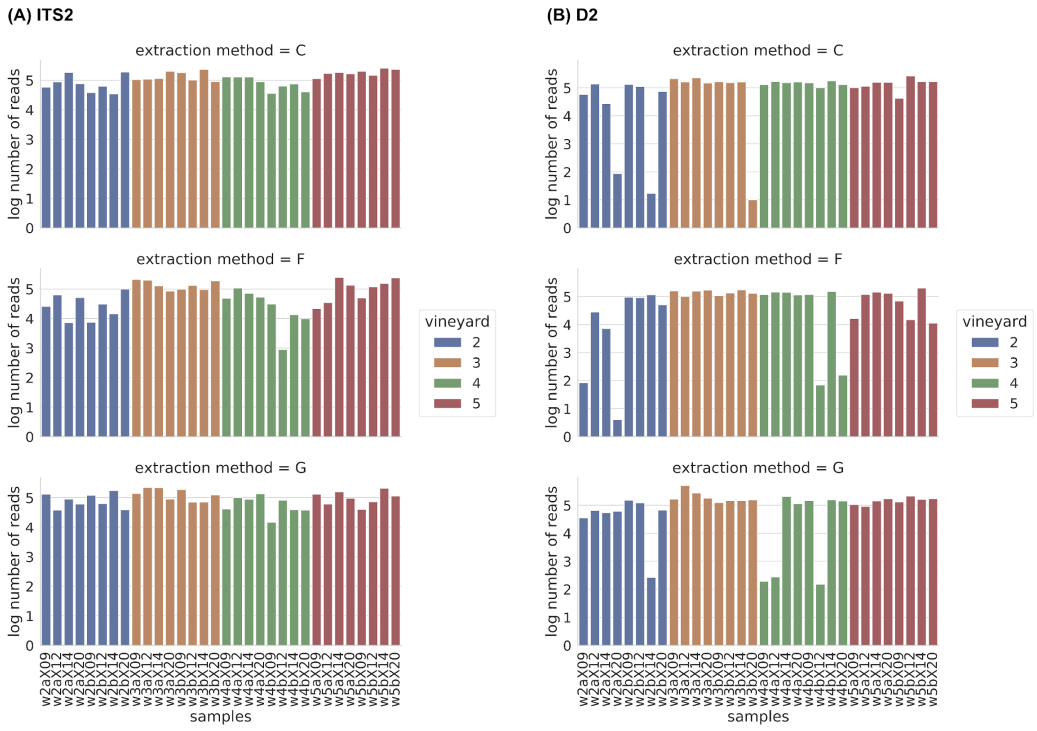


Fig. S3 The total number of read counts (in log 10 scale) of each extraction method for each sample separated by vineyard (2: blue; 3: orange; 4: green; 5: red) of (A) ITS2 and (B) D2 markers. Three methods were CTAB/PVPP-chloroform protocol (abbreviated as C: top),

commercial kit (F: middle) and G2 DNA/RNA enhancer (G: bottom). Samples with < 1000 reads (i.e. value 3 in the y-axis of plot) were removed from further analysis (ITS2 = 1 sample; D2 = 11 samples).

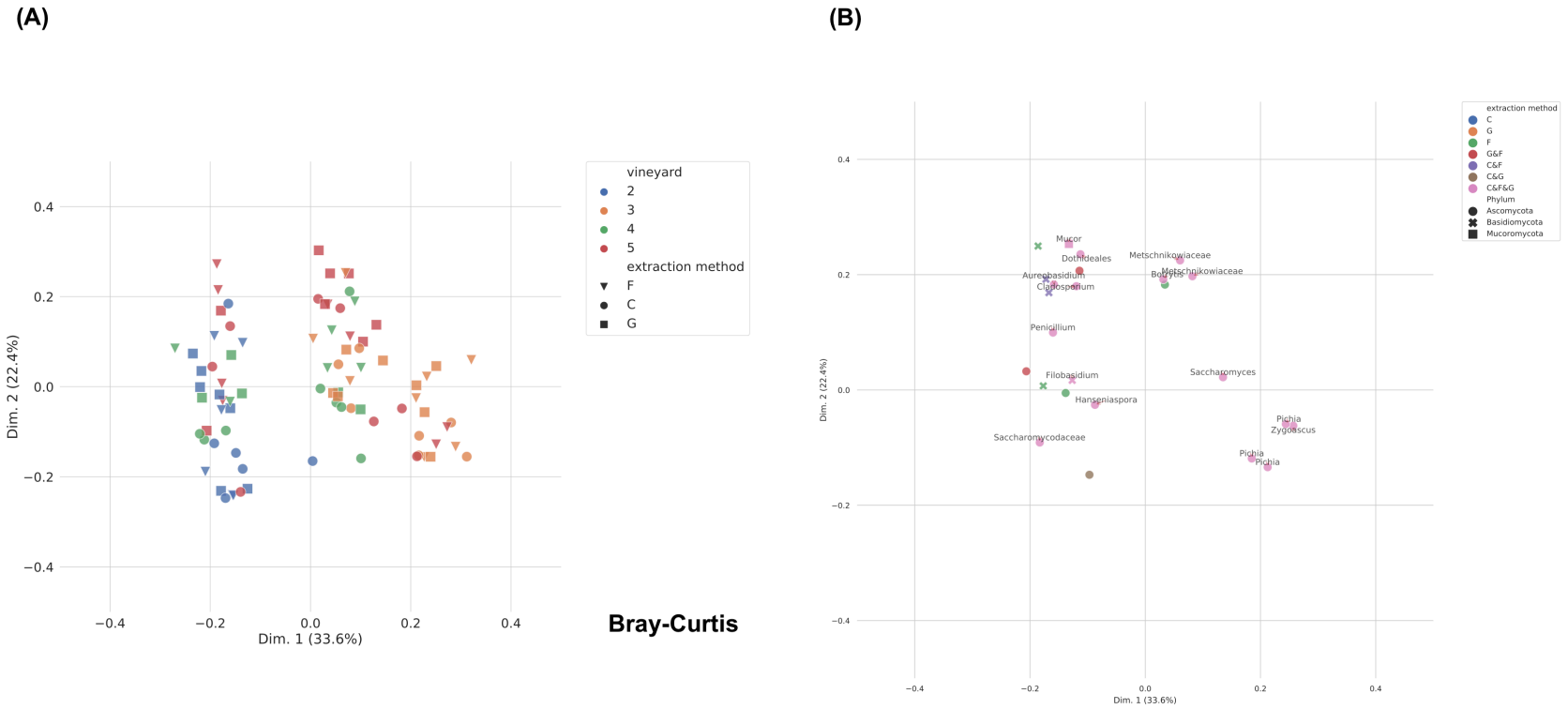


Fig.S4A & B β-diversity of fungal communities (left panel) and the corresponding variable loadings (right panel) from D2. PCoA plots from β-diversity of Bray-Curtis dissimilarity separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

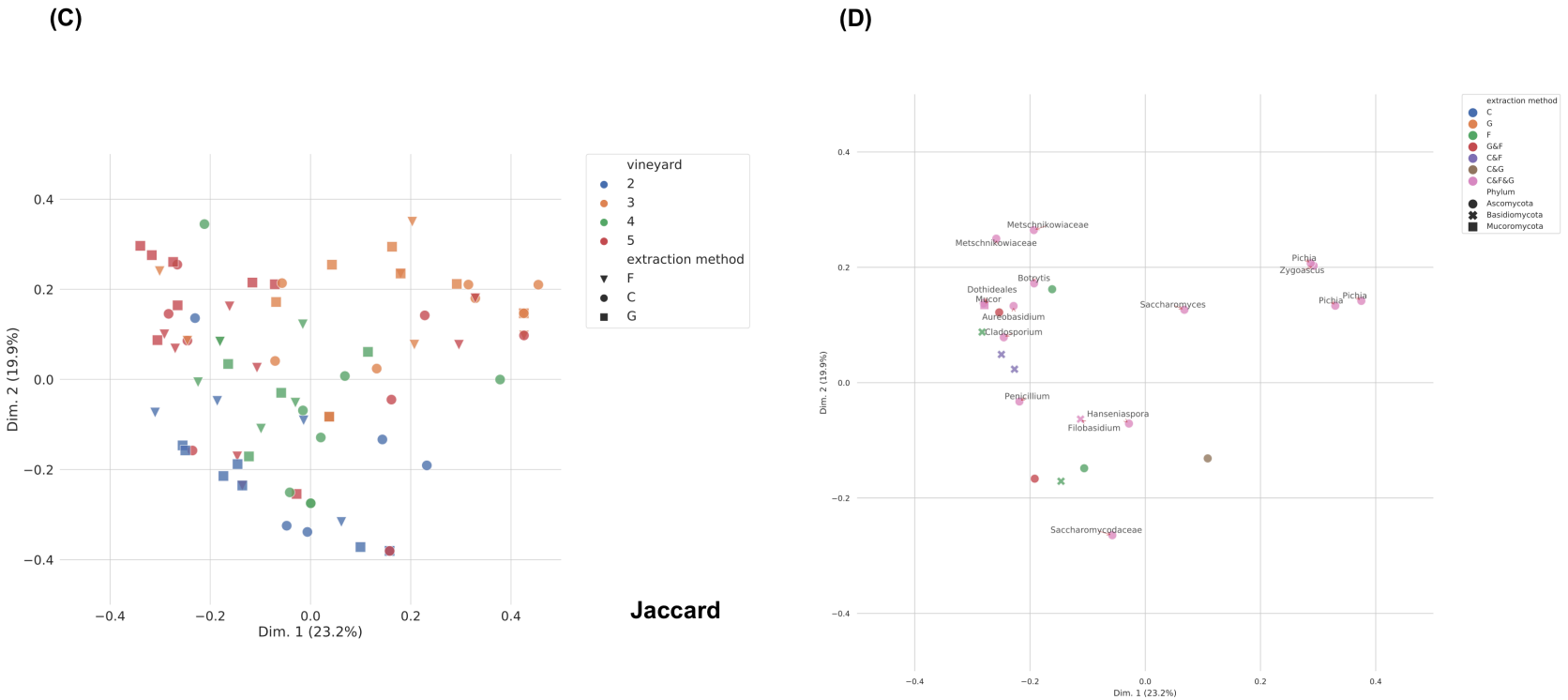


Fig.S4C & D β-diversity of fungal communities (left panel) and the corresponding variable loadings (right panel) from D2. PCoA plots from β-diversity of Jaccard index separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

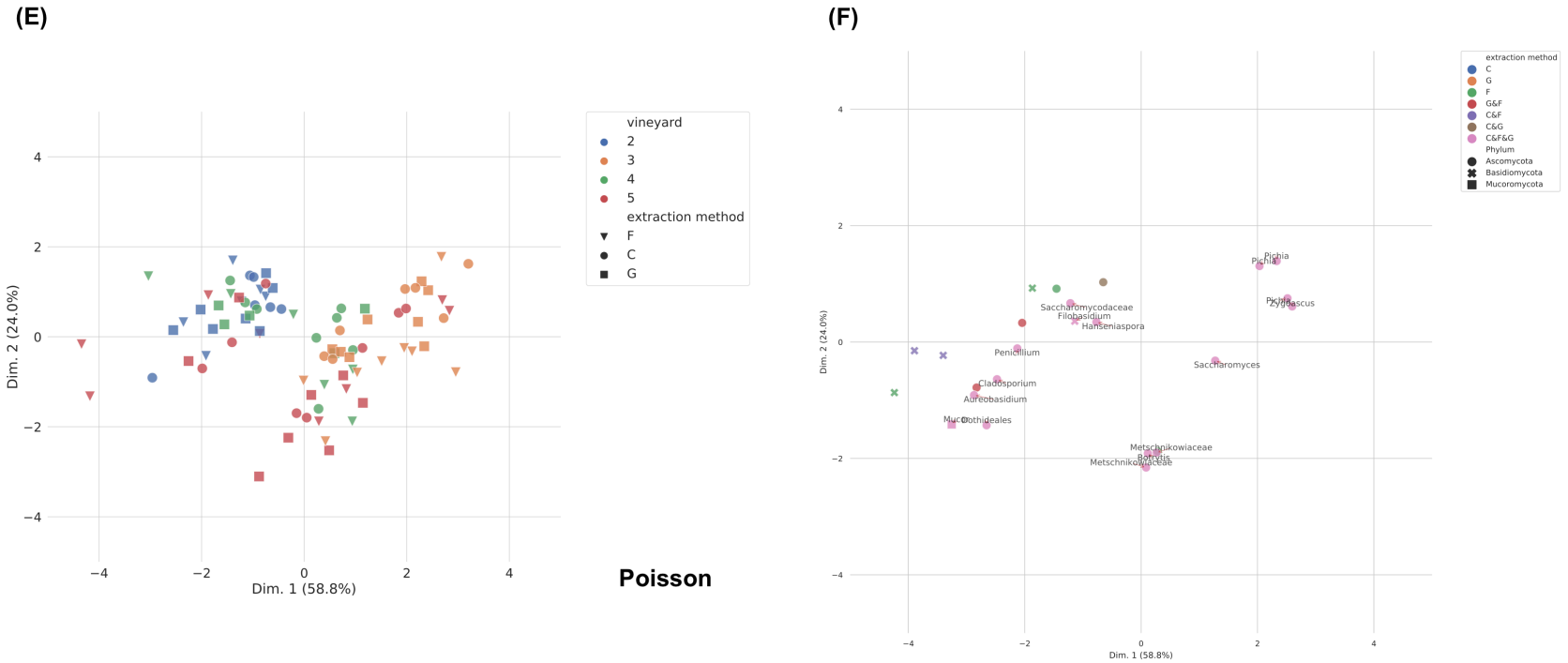


Fig.S4E & F β-diversity of fungal communities (left panel) and the corresponding variable loadings (right panel) from D2. PCoA plots from β-diversity of Poisson dissimilarity separated by different extraction methods (F: triangle; C:square; G: circle) and different vineyards (2: blue; 3: orange; 4: green; 5: red). For the corresponding variable loadings, OTUs are colored based on different extraction methods combinations they are found (C: blue; G: orange; F: green; G&F: red; C&F: purple; G&C: brown; all 3 methods: pink). Symbols correspond to different phyla (Ascomycota: circle; Basidiomycota: cross; Mucoromycota: square). OTUs commonly found in all extraction methods were labelled.

Chapter 4

Metagenomic and functional analysis of microbial composition in Riesling during spontaneous in vitro fermentation

Taxonomic and functional characterisation of the microbial community during spontaneous in vitro fermentation of Riesling must

Kimmo Sirén[1,2†], Sarah Siu Tze Mak[3†], Chrats Melkonian[4], Christian Carøe[3], Jan Hendrik Swiegers[5], Douwe Molenaar[4], Ulrich Fischer[1], M. Thomas P. Gilbert[3,6*]

[1]Institute for Viticulture and Oenology, Dienstleistungszentrum Ländlicher Raum (DLR) Rheinpfalz, Neustadt an der Weinstraße, Germany

[2]Department of Chemistry, University of Kaiserslautern, Kaiserslautern, Germany

[3]Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

[4]Systems Bioinformatics, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

[5]Chr. Hansen A/S, Boege Alle, 2970, Hoersholm, Denmark

[6]NTNU University Museum, Trondheim, Norway

[†]**authors contributed equally**

*** Correspondence:**
Prof. Dr. M. Thomas P. Gilbert
tgilbert@snm.ku.dk

Keywords: shotgun sequencing[1], metabarcoding[2], wine[3], microbial diversity[4], alcoholic fermentation[5].

Manuscript: 6853 words

Figures in manuscript: 4

Tables in manuscript: 1

Figures & Tables for submission (including Supplementary information): 15

Abstract

Although there is an extensive tradition of research into the microbes that underlie the winemaking process, much remains to be learnt. We combined the high-throughput sequencing (HTS) tools of metabarcoding and metagenomics, to characterise how microbial communities of Riesling musts sampled at four different vineyards, and their subsequent spontaneously fermented derivatives, vary. We specifically explored community variation related to three points: i) how microbial communities vary by vineyard; ii) how community biodiversity changes during alcoholic fermentation; and (iii) how microbial community varies in musts that successfully complete alcoholic fermentation versus those that become 'stuck' in the process.

At the vineyard level, our metabarcoding data shows an overall impact, and difference in the abundance of *Metschnikowia* between vineyards 4 and 5. We also observed at these two vineyards an

increase in non-Saccharomycetaceae fungal functions and a decrease in bacterial functions during the early fermentation stage. To investigate this further, we reconstructed draft genomes (bins), yielding 11 coherent bins and found that these vineyards shared yeast bins *Hanseniaspora* and *Saccharomyces*. Read recruitment and functional analysis revealed that during fermentation, a high abundance of *Metschnikowia* might serve as a biocontrol agent against bacteria, via a putative iron depletion pathway, and this in turn could help *Saccharomyces* dominate the fermentation.

With regards to alcoholic fermentation, we observed a general decrease in biodiversity in both the metabarcoding and metagenomic data. Unexpected *Micrococcus* behaviour was observed in vineyard 4 according to metagenomic analyses based on reference based read mapping. Analysis of open reading frames using this data showed an increase of functions assigned to unknown Actinobacteria in the end of fermentation. Therefore, we hypothesise that bacteria might sit-and-wait until *Saccharomyces* activity slows down.

Lastly, our metabarcoding data enabled us to identify a relationship between stuck fermentations and *Starmerella* abundance. Given that chemical analysis indicated that although the stuck samples contained residual glucose, all fructose had been consumed, we hypothesise that fructophilic *Starmerella* dominated these fermentations instead of *Saccharomyces*. Overall our results showcase how metagenomic functional analysis offer possibilities to improve our insights into the wine alcoholic fermentation process, including highlighting microbe interactions.

1 Introduction

Microbial interactions are vital to the winemaking process, with numerous different microbes known to be involved in the formation of wine flavour and aroma. Understanding this microbial diversity and their interactions throughout the process stands to enhance our knowledge of winemaking and wine complexity (Tempère et al., 2018). Although research on wine microbes has a long history (Pasteur, 1872), many significant challenges remain to be solved, not least due to difficulties in studying the composition of wine's complex matrix. In this regard there is considerable interest in the application of High-Throughput Sequencing (HTS) tools such as metabarcoding and shotgun metagenomic sequencing to wine research, given their potential to offer us more in-depth characterisation of the microbial community (Belda et al., 2017; Sirén et al., 2019; Stefanini and Cavalieri, 2018).

Key questions that have received considerable attention include the origin of yeasts that drive the fermentation, and how different microbes shape the fermentation (Bisson et al., 2017; Fleet, 1990, 2003; Pretorius, 2000; Varela and Borneman, 2017). The answers to these questions are not clear cut. For example, even whether sufficient *Saccharomyces cerevisiae* is present in the vineyard (thus entering the must during pressing) to drive fermentation is debated (Martini, 1993). This question is particularly timely today, given the current trend to return to spontaneous fermentation during winemaking, for reasons relating to both typicality as well as arguments that spontaneously fermented wines gain in complexity due to the more diverse microbial interactions (Di Maro et al., 2007). Furthermore, how the vineyard and winery flora actually interacts remains inconclusive. While some authors have suggested that the main contributors to fermentation originate from the vineyard flora (Bokulich et al., 2014, 2016; Morrison-Whittle and Goddard, 2018), others argue that the winery dominates (Ganucci et al., 2018; Stefanini et al., 2016). In short, both the individual effects, and the interaction of the two continue to remain elusive.

A further topic of interest is the dynamics of the microbial community during the alcoholic fermentation. While alcoholic fermentation is known to result from a succession of various microbes, with *Saccharomyces* eventually dominating, details about the timing and abundances of different microbes remain of interest (Stefanini et al., 2016). It is currently understood that while microbial

diversity decreases during the winemaking process (Bisson et al., 2017), some microbes can survive (Romano et al., 2003; Torija et al., 2001), such as *Metschnikowia pulcherrima* (Díaz et al., 2018) and bacteria (Bokulich et al., 2015; Piao et al., 2015). Although traditionally how the different microbial species interact has been studied using culture-based techniques, they are increasingly targeted using culture-independent methods (Bagheri et al., 2017; Jolly et al., 2014; Morgan et al., 2018; Sternes et al., 2017; Zepeda-Mendoza et al., 2018).

Another area of growing interest relates to the role of spontaneous fermentation. Because this is driven principally by non-*Saccharomyces* yeasts, spontaneous fermentations are regarded as being able to diversify aromatic quality (Liu et al., 2016; Zott et al., 2011). However, their use remains intimidating for the industry, as they can lead to unwanted characteristics (Ciani et al., 2006), and sluggish, or even stuck, fermentations (Bisson 1999; Bisson and Butzke 2000). While the main reason for sluggish fermentation is often nutrient related, microbial interaction could play a role. Recently, it has been suggested based on metabarcoding data that high species richness including the presence of non-*Saccharomyces* yeasts in must samples can relate inversely to the capability of *Saccharomyces* to carry out the fermentation (Boynton and Greig, 2016). Therefore, adding sulphur (SO2) to the harvested grapes or must might also help to remove competition originating from the vineyard microbes. This might help to advance the spontaneous fermentations by reducing the competition by allowing the winery microbes to become dominating.

Although there is an increasing trend to apply HTS tools to studies of wine microbiology, with few exceptions these have been amplicon-based 'metabarcoding' approaches that enable community profiling (e.g. as varies by geography, regions, *Botrytis*) as reviewed extensively by Belda et al. (Belda et al., 2017) and Stefanini and Cavalier (Stefanini and Cavalieri, 2018). Thus, there is considerable interest in profiling the whole genomic content using true metagenomic approaches, in the hope that they will also provide information about the functional pathways involved (Sirén et al., 2019). In this regard, there are currently only two shotgun sequencing based metagenomic studies of wine fermentation published (Sternes et al., 2017; Zepeda-Mendoza et al., 2018). The first investigates the bias between metagenomics and metabarcoding in assessing community structures while the second one characterises the impact of the inoculations of *Oenococcus oeni* and *Brettanomyces bruxellensis* on volatile phenol formation.

Given the potential of HTS for investigating microbial community and interactions, and metagenomics in providing insights into genomes and genes together with their functional potential (Sirén et al., 2019), in this study we combined metabarcoding and metagenomic sequencing in order to investigate the microbial community differences and biodiversity decrease during fermentation when the fermentation is solely carried out by vineyard microbes. We chose to spontaneously ferment must samples originating from four Riesling vineyards in Pfalz (Germany). While most of the samples completed alcoholic fermentation, but several exhibited sluggish fermentation. This enabled also enabled us to examine the microbial basis of this phenomenon.

2 Material and Methods

2.1 Fermentation set up and sample collection

A fermentation experiment was carried out with grapes from four Riesling vineyards in the Pfalz wine region of Germany in Autumn 2015, in order to investigate spontaneous fermentation dynamics, i.e. fermentation derived solely from vineyard, rather than winery, microbes. The grapes were handpicked into autoclaved sterile flat plastic bags (Neolab) which were sealed inside the vineyards and processed at the Institute for Viticulture and Oenologie, Dienstleistungszentrum Ländlicher Raum Rheinpfalz, Neustadt an der Weinstraße, Germany. The microvinifications were carried out in sterile conditions under a laminar flow hood, to restrict fermentation to only those

microbes present in the vineyards. After crushing and pressing, the must from each vineyard was left in 3 L sterile autoclaved Erlenmeyer flasks (Duran, Germany) for overnight sedimentation at 4 °C. After settling, must was racked into duplicate 1 L autoclaved Erlenmeyer flasks (Duran, Germany) and secured with silicone bungs attached to distilled water filled airlocks. Sample collection was done over four different days during alcoholic fermentation, with 4.5 mL of fermenting must collected at each timepoint for subsequent DNA analysis (total 32 extracts, Table 1), and 40 mL collected for monitoring of the fermentation by measuring various wine parameters (alcohol, density, total sugar, glucose, fructose, glycerol, titratable acidity, pH, tartaric acid, malic acid, lactic acid, citric acid, volatile acid, glycerol, yeast assimilable nitrogen, primary amino nitrogen (NOPA) and ammonium with routine FTIR analysis (WineScan FT120, FOSS Electric). Furthermore, the alcohol percentage was estimated by dividing the product of measured alcohol (g/mL) and sample density (g/mL) by 10 times the density of ethanol (g/mL). In general, estimation of the fermentation progress through changes in density or alcohol concentration is complicated without continual monitoring of the must. However it has been observed that continuous monitoring can potentially expose must to contaminating microbes (Grangeteau et al., 2015; Ocón et al., 2013; Pérez-Martín et al., 2014), and therefore, we chose instead to sample four times during the fermentations. Details of samples and measured wine parameters are shown in Supplementary Table 1.

2.2 DNA extraction

Prior to DNA extraction, each sample, was centrifuged at 4500×g for 10 minutes, after which the supernatant was removed and the pellet resuspended with 1 mL of ice-cold 1X PBS (pH 7.4, Life Technologies, California, U.S.A.). The resuspended samples were washed twice with 1 mL of ice-cold 1X PBS to remove debris. The pellets were subsequently stored at -20 °C until DNA extraction. DNA extractions were performed as described in a parallel study (Mak et al., in review), with the use of FastDNA Spin Kit for Soil (MP Biomedical, California, U.S.A.) following the manufacturer's protocol with minor modifications. In brief, pellets were bead-beaten twice at 30 Hz for 40 seconds using a TissueLyser II (Qiagen, Hilden, Germany), with cooling step on ice for 2 minutes in between bead-beating steps. In the elution step, 105 µL of 1X TET buffer (1X TE buffer in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0, Sigma-Aldrich and 0.05% Tween 20, Sigma-Aldrich) was added to the filter column then incubated at 55 °C for 5 minutes before elution. DNA was subsequently subjected to an extra purification step using a DNA Clean and Concentrator™-5 (Zymo Research, California, U.S.), and eluted in a final volume of 55 µL of 1X TET buffer. DNA extracts were quantified using a Qubit 1.0 fluorometer with dsDNA High Sensitivity Assay kit (ThermoFisher Scientific). An extraction blank was included for every 16 samples.

2.3 qPCR

Prior to metabarcoding PCRs, we used quantitative real time PCR (qPCR) to both estimate the number of copies of the region, thus determine the number of PCR cycles, and to identify whether PCR inhibitors were present in the DNA extracts. For both qPCR and metabarcoding we used fusion primers targeting the fungal internal transcribed spacer 2 region (ITS2, ITS7_F from Ihrmark et al., 2012 and ITS4_R from White et al., 1990), each containing an exclusive 8 bp multiplex identifier tag (MID tag) and MiSeq sequencing adapters.

Each qPCR reaction consisted of a 25 µL reaction volume containing 2 µL of template and 23 µL of mastermix containing 1X GeneAmp® 10X PCR Buffer II (Applied Biosystems, U.S.A.), 2.5 mM $MgCl_2$ (Applied Biosystems, U.S.A.), 0.8 mg/ mL Bovine Serum Albumin (BSA), 1 µL SYBR Green (Invitrogen, CA, USA), 0.25 mM dNTPs, 0.4 µM forward primer, 0.4 µM reverse primer, 0.25 µL AmpliTaq Gold DNA polymerase (Applied Biosystems, U.S.A.) and 14.5 µL AccuGene molecular biology water (Lonza). qPCR conditions were as follows: 95 °C for 5 minutes, followed by (95 °C for 30 seconds, 52 °C for 30 seconds, and 72 °C for 45 seconds) for 45 cycles, and a final

dissociation curve of 1 cycle of 95 °C for 1 minutes, 55 °C for 30 seconds and 95 °C for 30 seconds. PCRs were performed using a MX3005 qPCR machine (Agilent). Standard curves with duplicates made by a serial dilution of PCR products. Positive controls and negative controls were included.

2.4 PCR and metabarcode sequencing

We applied metabarcoding to all DNA extracts and controls. Metabarcoding PCRs were based on the same mastermix as that used in the qPCR, except replacing 1 µL SYBR Green (Invitrogen, CA, USA) with 1 µL AccuGene molecular biology water (Lonza, Switzerland). PCRs were carried out in AB 2720 Thermal cycler (Applied Biosystems, U.S.A.) with the following conditions: 95 °C for 5 minutes, followed by 33 cycles of 95 °C for 30 seconds, 52 °C for 30 seconds, and 72 °C for 45 seconds, and a final elongation step of 72 °C for 10 minutes. Extraction blanks, positive controls and PCR negative controls were included. PCR products (~480 bp) were visualised by electrophoresis using 2% agarose gels, then subsequently pooled together with amplicons derived from a parallel study (Mak et al., in review) into 3 amplicon pools. Amplicon pools were subsequently purified with QiaQuick columns (Qiagen) following the manufacturer's protocol to remove primer dimers. An aliquot of each amplicon pool was used for quantification and size estimation using the High-Sensitivity D1000 Screen Tape for Agilent 2200 TapeStation (Agilent). Lastly, purified amplicon pools were sent for sequencing in 2 flow cells on the Illumina MiSeq platform in 250 bp paired-end mode at The Danish National High-Throughput DNA Sequencing Centre, Copenhagen, Denmark. This dataset comprised of ⅔ of a MiSeq flow cell.

2.5 Shotgun library construction and metagenomics sequencing

We additionally generated shotgun metagenome data on a subset of ten samples. These were chosen as samples exhibiting normal fermentation rate, with a focus on varying alcoholic levels different vineyards, and given the fungal diversity observed during the metabarcoding. Metagenomic sequencing (Table 1) was performed with BGISeq technology (Fang et al., 2018), although with a customised library build. The DNA extracts were initially fragmented to around 300 bp using a Bioruptor 300 (Diagenode, Belgium) using 10 cycles of 30s on and 90s off. The DNA was then converted into indexed sequencing libraries using the Blunt End Multi Tubes (BEMT) protocol (Supplementary File 1, Supplementary Figure 5) following an initial comparison of the performance of different library construction methods on DNA extracted from ferment samples (Supplementary File 1). Library blanks and index PCR blanks were included. Thirty µL (input amount: < 0.3 - 6.6 ng) of each sample was used for each library constructions (Supplementary Table 2) and 2 µL of 10 µM BGI 2.0 adapters (Supplementary Table 3) were added to each sample in the adapter-ligation step. Each library was quantified using qPCR post construction in order to determine the appropriate number of PCR cycles to subject each to. This was done using a MX3005 qPCR machine (Agilent) with the forward primer and one indexed reverse primer. Post qPCR, each library was subsequently PCR amplified and indexed with different indices and purified for residual adapter dimers using SPRI beads (Sigma-Aldrich) in 1.5X beads:library ratio with incubation at 37 °C for 10 min and eluted in 50 µL. All libraries were then sent to BGI-Europe for sequencing, where they were pooled in equimolar concentrations for circularisation, DNA nanoballs (DNB) construction and sequencing on the BGISeq-500 platform over 4 lanes in 100bp pair-ended mode.

2.6 Sequence data analyses

2.6.1 Metabarcoding sequencing analysis

Metabarcoding sequence analyses were performed following the pipeline described in Feld et al. (2016), although with modifications in trimming, post-clustering and the use of databases. Raw reads were merged and demultiplexed using vsearch v2.1.2 (Rognes et al., 2016). Cutadapt v1.11 (Martin, 2011) was used for removal of adapters and primers. Reads smaller than 100bp were trimmed with

vsearch, followed by dereplication. Singletons and chimeras were filtered using the UPASE pipeline (Edgar, 2013), and the reads were clustered to operational taxonomic units (OTUs) with the command -cluster_otus. Reads were mapped back (including singletons) to the filtered clusters with 99% similarity using usearch v9.0.2132 (Edgar, 2010) in order to create the OTU table, then subjected to the post-clustering algorithm LULU (Frøslev et al., 2017) as implemented in R v3.4.1. Filtered OTUs were then aligned with the reference UNITE+INSD database released on 2017.12.01 (UNITE Community 2017) for taxonomic assignment to genus level, with 97% identity threshold, 70% coverage in BLAST using QIIME v1.9.1 (Caporaso et al., 2010) with a modified assign_taxonomy.py script. OTUs that did not obtain taxonomic assignment in the above were labelled as "No blast hit".

Metabarcoding data were analysed using *phyloseq* (McMurdie and Holmes, 2013) framework in R (v3.4.0). OTUs with fewer than 10 reads (Oliver et al., 2015; Werner et al., 2012) and samples with fewer than 1000 reads were discarded for further analyses. Furthermore OTUs with "No Blast Hit" were also removed from analyses. The α-diversity was evaluated with relative abundances and Hill numbers (Hill, 1973). Hill numbers were calculated using q-values between 0 and 3 with 0.001 intervals (Alberdi and Gilbert, in review). The resulting matrix was decomposed with Principal Components Analysis (PCA), with retention of the first principal component as a measure of the overall effect of α-diversity. The variance of the Hill numbers was additionally retained for visualisation. Heatmap visualisation and clustering were done on variance stabilised transformed (Anders and Huber, 2010) count data using *DESeq2* (Love et al., 2014) with hierarchical clustering using weighted linkage (WPGMA) (Sokal and Michener, 1958) on Pearson correlation. Furthermore differential abundances and corresponding log2 fold changes and adjusted p-values were found using *DESeq2* (Love et al., 2014). *DESeq* function parameters were set as following: test type: 'Wald'; fittype: 'parametric'. Significant differences between relative abundances were controlled by setting false discovery rate (FDR) at 5% using the method by Benjamini and Hochberg (Benjamini and Hochberg, 1995). OTUs whose abundances differed significantly between vineyards were visualised with letter-value boxplots (boxenplots) (Hofmann et al., 2017) of the raw count data.

2.6.2 Metagenomic data analyses

The metagenomic data analysis consisted of two stages. In the first stage, the raw sequence reads were analysed individually for each sample. In the second stage, a binning approach was used, where analyses were performed on the total sequence data from each of the two vineyards (specifically vineyard 4 and vineyard 5). This was used to specifically explore for differences in the microbial communities relating to alcoholic fermentation or vineyard of origin.

For all analyses, the qualities of all paired-end reads were checked using FastQC (Andrews, 2010), both before and after adapter removal using Trim Galore (Krueger, 2012) and Cutadapt (Martin, 2011), with the following parameters: default Phred score: 20 and cutoff for read length: 40 bp for filtering. All reads containing 'N' were also filtered and the filtered reads were subsequently merged to compile with IDBA-UD (Peng et al., 2012) for sequence assembly. MEGAHIT v1.0.4 (Li et al., 2016) was then used for sequence assembly, with metagenomic parametrisation for either individual samples (Stage 1) or binning by individual vineyards (Stage 2).

For taxonomic assignment, a curated database was constructed using the genomes of 130 relevant species, including eukaryotes and bacteria. This database is similar to, but expanded on, that used in a previous study (Sternes et al., 2017) (Supplementary Table 4b). These sequences were obtained from NCBI and used to construct a custom Kraken database. Relative abundances were obtained by mapping using Kraken v.1.0 (Wood and Salzberg, 2014) and Braken (Lu et al., 2017). Further data analysis was performed in R v.3.4.4 with *phyloseq* (McMurdie and Holmes, 2013) and other custom scripts. The Kraken mapped reads that were higher than 0.00001 relative proportion were retained for

clustering which was done on Pearson correlation similarities using affinity propagation (Bodenhofer et al., 2011; Frey and Dueck, 2007). A general workflow to assess the most suitable number of clusters was started by setting the exemplar preferences value high, which led to a very large number of clusters. Application of agglomerative clustering on the resulting affinity propagation clusters, allowed an inspection of the corresponding dendrogram. Based on the dendrogram, a cutoff was manually decided and affinity propagation was rerun repeatedly to achieve the desirable number of clusters. Prodigal (Hyatt et al., 2012) was used to predict the open reading frames (ORFs), with parameterisation for metagenomes of individual samples (Stage 1), or binned by samples per vineyard (Stage 2), as detailed above. In order to perform the functional analysis, the ORFs were then used as an input for eggNOG (Powell et al., 2014), to obtain KO (KEGG Orthology) assignments and Clusters of Orthologous Groups (COGs) as functional annotations. The unique KOs generated for each sample individually were combined using custom Python and R scripts and were used to to create a matrix of total 8744 KO, which were combined in a matrix with unique 412 pathways according to KEGG for all samples together. The KO pathway enrichment analysis was done using the following packages: Biostrings (Pagès et al., 2017), ggplot2 (Wickham, 2016), reshape (Wickham, 2007), KEGGREST (Tenenbaum, 2018), lattice (Sarkar, 2008), apcluster (Bodenhofer et al., 2011; Frey and Dueck, 2007), BioServices (Cokelaer et al., 2013) and pandas (McKinney, 2011).

For the binning approach (Stage 2), assemblies were performed with MEGAHIT on the pooled data from vineyards 4 and 5 respectively, in order to create their corresponding contig files. Next, the trimmed sequences from the data pre-processing steps for each sample were mapped to their corresponding vineyard assembled contigs files using the bwa-mem algorithm v0.7.15 (Li, 2013). The mapped sequences were subsequently cleaned of PCR duplicates using samtool v1.6 (Li et al., 2009) and exported as BAM files for binning. Binning was subsequently performed and the output was visualised with the metagenomics workflow (http://merenlab.org) in Anvi'o v5.1's interactive interface (Murat Eren et al., 2015). Each assembled contig was used to create corresponding Anvi'o contig databases using the default settings. The databases were then run under HMMER v3.1.b2 (HMMER, 2015) for sequence searching using hidden Markov models (HMMs), and genes were annotated with functions from the NCBI's Clusters of Orthologous Groups (COGs) (Tatusov et al., 2003) using command anvi-run-ncbi-cogs. The genes in the contig database were classified using Kaiju v1.5.0 (Menzel et al., 2016), with NCBI's non-redundant protein database including fungi and microbial eukaryotes. Each of the sample bam files derived from a single vineyard was profiled with their corresponding annotated contigs database, with minimum contig length set to 2500 nt using the command anvi-profile, then merged to generate bins using CONCOCT as implemented in Anvi'o. Bins with completeness ≥ 40% and redundancy ≤ 10% were retained for analysis (Delmont et al., 2015). Abundance of the reads was estimated in terms of percentage of read recruitment. This was calculated by the mean coverage of each split in each bin with normalisation of all bins respect to each other in each sample. For functional assignment of the bins, 6394 KOs were generated when combining the 11 bins, which resulted in 411 unique pathways.

All figures, except those from Anvi'o were visualised using ggplot2 (Wickham, 2016), Matplotlib (Hunter, 2007) and Seaborn (Waskom et al., 2017), with further post-processing done in Inkscape v0.91 (https://inkscape.org/). The sequencing data were deposited to European Nucleotide Archive under study number: PRJEB30801 and ERS3017411-ERS3017414, ERS3017423-26, ERS3017435-38, ERS3017447-50, ERS3017459-62, ERS3017471-74, ERS3017483-86, ERS3017495-98 in study number: PRJEB29796.

3 Results

Eight different Riesling musts containing four vineyard specific microbial compositions were allowed to ferment spontaneously. While most of the samples were found to contain sufficient microbes for performing alcoholic fermentation (Supplementary Table 1), three exhibited sluggish

(henceforth referred to as 'stuck') fermentation behaviour. Specifically, these needed 6 months (data not shown) to finish the fermentation, as opposed to the 5 weeks time needed by the others (Figure 1A). Furthermore, no malolactic fermentation was observed (Supplementary Table 1).

Metabarcoding was performed by amplifying the ITS2 gene for the 32 extracts (four sampling dates for the 8 samples). Prior to subsequent analysis, data from one extract (sample w4b112 representing the stuck fermentation behaviour group) was removed due to yielding fewer than 1000 reads. In total, 2.79 million reads were generated yielding 105 OTUs, although after all filtering this was reduced to 2.75 millions reads representing 72 OTUs that were retained for subsequent analyses (Supplementary Table 5).

A total of 1.6 billion raw reads was generated from the 10 shotgun metagenomic sequenced samples, of which 1.5 billion were retained after adapter removal. Details of the number of reads per sample and the percentage that mapped to the curated database are shown in Supplementary Table 4.

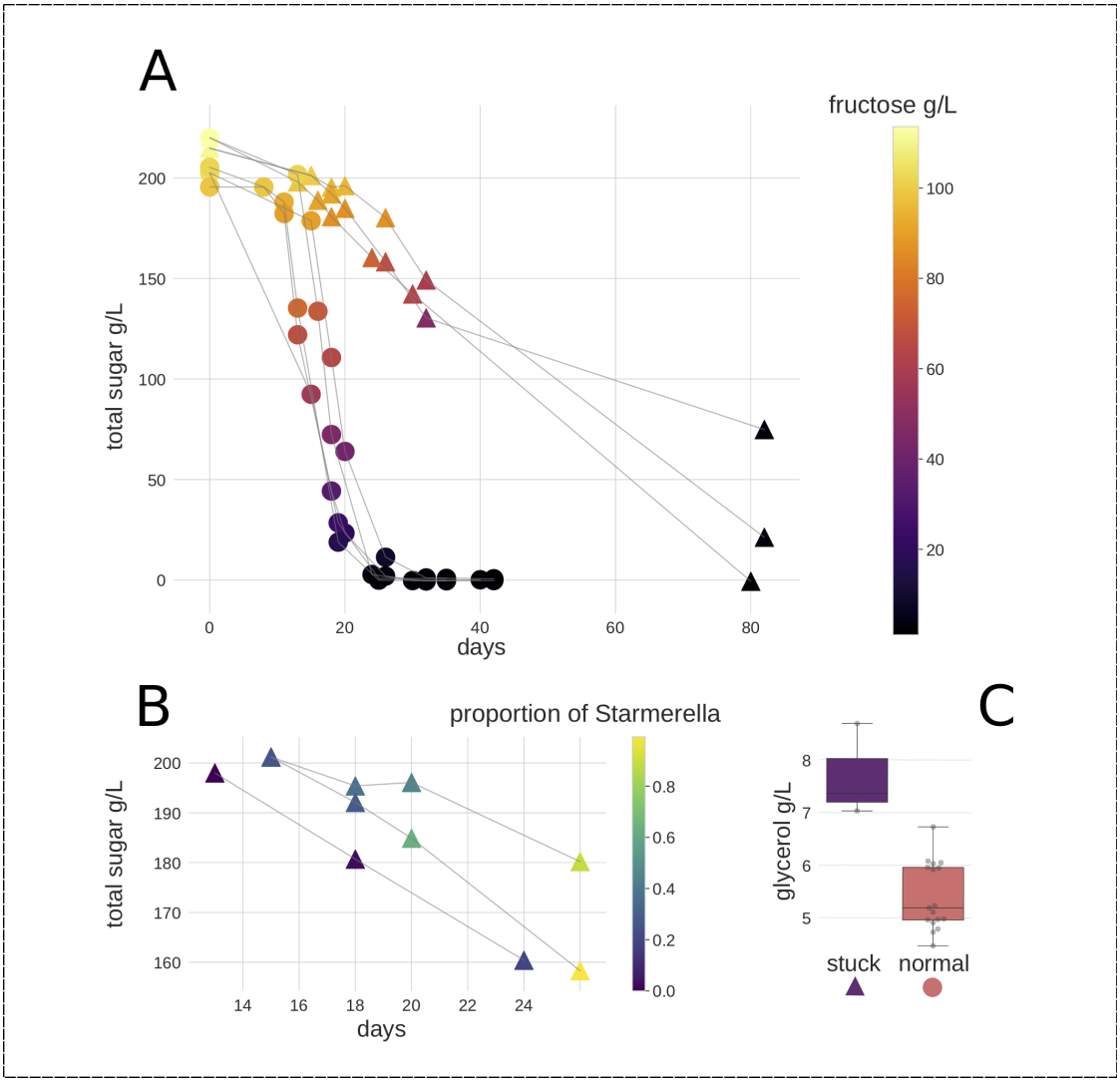


Figure 1. Fermentation performance of the samples. Triangles represent stuck behaviour samples and circles represent normal behaviour samples. **(A)** Most samples showed normal alcoholic fermentation, three samples stood out with a slower fermentation speed, and showed a preference for fructose instead of glucose. **(B)** A zoom in of stuck behaviour samples between fermentation days 13-26. *Starmerella* was found to be the highest in relative abundance in these samples in ITS2 region. **(C)** The glycerol concentration was higher in these stuck behaviour samples.

3.1 Overall community differences

Overall, while the results clearly show that *Saccharomyces* drives the alcoholic fermentation (given their abundance among the data), we observed three main drivers of sample clustering that relate to the microbial composition. Firstly, there were differences in fermentation behaviours; secondly, there were differences between the four vineyards; and thirdly there were clear differences relating to the stage of fermentation as expressed as alcohol percentage (Figure 2). Using these results, we subsequently explored the following questions. Firstly, we used both our metabarcoding and

metagenome data to explore how the microbial communities vary between vineyards. Secondly, we investigated how the stage of alcohol fermentation impacts the microbial biodiversity. This was done using both the metabarcoding data from all the samples, as well as the shotgun metagenomic data generated from several samples selected to explore the longitudinal effect at two of the vineyards. Thirdly, we used the metabarcoding data to investigate how the microbial community profile differs between the normal and stuck fermentations.

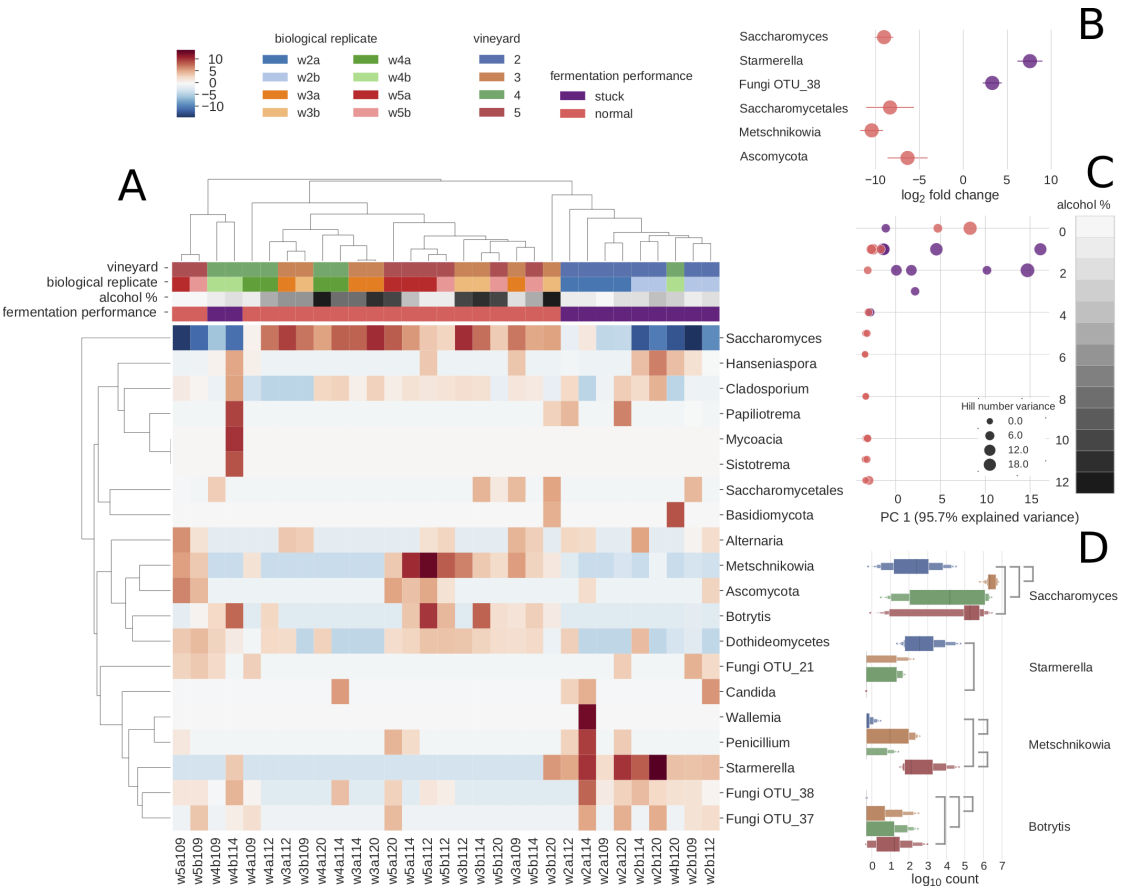


Figure 2. Microbial diversity explains sample differences using ITS2 metabarcoding. **(A)** Clustered heatmap depicted with 20 most abundant species. Species shown with Poisson dissimilarities. Clustering is mainly driven by fermentation performance. **(B)** Differential abundance of OTUs in log2 fold change by the fermentation performance: normal fermentation (pink) and stuck fermentation (purple). Four OTUs are found to be significantly more abundantly expressed in normal fermentations, whereas two OTUs are found significantly more abundantly expressed in stuck fermentations. **(C)** The relationship of alcohol percentage and α-diversity was measured as a first principal component of decomposed Hill numbers with q-values between 0 and 3, for all samples is found to decrease and vary more with rising alcohol percentage. The size of dots increases with the increase in Hill's number variance. **(D)** Vineyards were studied with pairwise comparisons of differentially expressed abundances. Line between vineyards indicates a significant difference between the pairwise comparison.

3.2 Differences between vineyards

The differences between the four vineyards were initially investigated by metabarcoding all samples for the ITS2 region. Pairwise comparisons of the vineyards indicated that vineyard 2 differed most from the three other vineyards, with the differential abundances of four fungal OTUs clearly standing out (Figure 2D). *Botrytis* (3: *p-adj*= 3.51e-7, 4: *p-adj*= 1.68e-3, 5: *p-adj*= 1.76e-5) and *Saccharomyces* (3: *p-adj*= 4.00e-8, 4: *p-adj*= 1.78e-4, 5: *p-adj*= 5.81e-5) were found to be significantly lower in this vineyard, whereas *Starmerella* was found at higher abundance in vineyard 2 than vineyard 5 (*p-adj*= 4.21e-06). Lastly, *Metschnikowia* was found to be significantly less abundant in vineyard 2 than vineyards 5 (*p-adj*= 1.39e-09) and 3 (*p-adj*= 7.56e-05). The only significant difference that was found when comparing the other vineyards was between vineyards 4 and 5 where one OTU, *Metschnikowia* (*p-adj*= 3.50e-4), was found significantly higher concentrations in vineyard 5 (Figure 1D).

These differences were further explored through metagenomic sequencing of the samples from vineyards 4 (W3 and W4) and 5 (W9 and W10), using samples taken at similar fermentation stages as estimated by alcohol percentage. Annotating reads to the curated taxonomy database showed that samples with lower alcoholic percentage had a lower percentage of mapped reads (W3: 15.33% and W9: 34.98%) compared to those with higher alcohol percentage (W4:77.69% and W10: 50.54%) (Supplementary Table 4a).

Next, a functional comparison of the vineyards was performed by using the COG classification derived from eggNOG annotation. The count number of bacterial functions was significantly higher in vineyard 4, whereas, the vineyard 5 had a small increase of Saccharomycetaceae functional count and a significant increase of the “other fungus” functional count (Supplementary Figure 1). In order to further explore this non-Saccharomycetaceae group, and to validate if this corresponded to the presence of *Metschnikowia* species (as observed with metabarcoding), binning was applied with the aim of reconstruction of draft genome.

Figure 3 shows an overview of the binning results and the comparison of the differences between vineyards 4 and 5 corresponding with regards to alcoholic levels. From vineyard 4, five genomic bins were obtained by assembly and binning. These varied in size between 1.87 Mbp and 9.78 Mbp, with completeness varying from 44.8-94.96%, and redundancy in 0-10% (Supplementary Table 6). For vineyard 5, six genomic bins were found. These varied in sizes between 2.34 Mbp and 12.3 Mbp, with completeness ranging 54.22-97.84% and redundancy in 0.72-9.35% (Supplementary Table 6). Interestingly, the vineyards had two yeast and two bacterial bins (*Hanseniaspora*, *Saccharomyces*, unknown Actinobacteria and *Pelomonas*) assigned to same taxa. Furthermore, vineyard 4 also had unique unknown bacterial bin while vineyard 5 was found to have two additional bins with one for non-*Saccharomyces* yeast, *Metschnikowia*, and another one for bacteria, *Bradyrhizobium* (Figure 2, Supplementary Table 6).

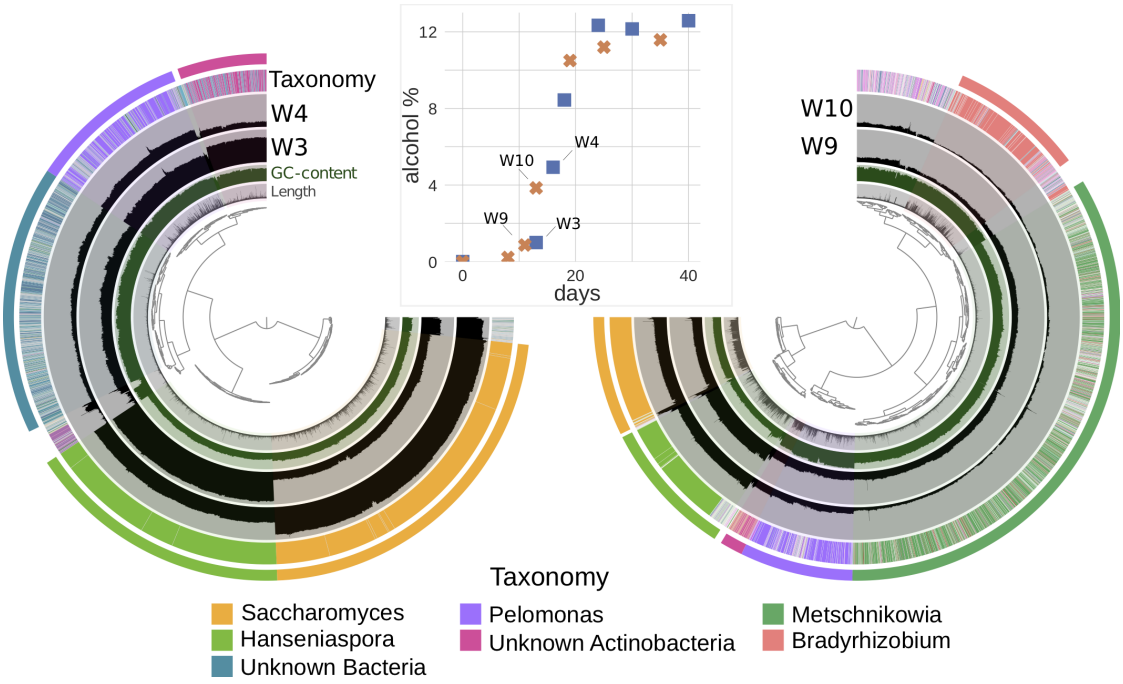


Figure 3. The visualisation of mean coverage of genome bins in samples W3 and W4 in vineyards 4 (left), and W9, W10 in vineyard 5 (right) using Anvi'o and their corresponding alcoholic percentages during fermentation (middle; blue square: vineyard 4; orange cross: vineyard 5). GC-content, length of each contig and taxonomy classified by Kaiju were displayed in adjunct layers. Each sample is represented in a separate layer, and each bar inside the sample layer corresponds to a datum computed for a given split, where contigs longer than 20k bp were divided to different splits. The outermost layer was the bin layer, where corresponding colours linked with the taxonomy of each bin. Vineyard 4 with 4333 contigs (minimum length: 2500 nucleotides, total nucleotides: 27.29 Mbp) that represented 4% of all contigs and 36% of all nucleotides found in the vineyard 4 contigs database (93546 contigs and total nucleotides: 74.84 Mbp). Vineyard 5 with 5364 contigs (minimum length: 2500 nucleotides, total nucleotides: 53.01 Mbp) that represented 5% of all contigs and 50% of all nucleotides found in the vineyard 5 contigs database (92948 contigs and total nucleotides: 105.44 Mb). The hierarchical clustering of contigs based on the sequence composition and their sample distribution were used for the dendrograms at the centre of Anvi'o visualisation. More bin details are shown in Supplementary Table 6.

As expected in the functional analysis, differentiation was found between fungi and bacteria in the form of 2 distinct clusters (Supplementary Figure 2). Moreover, for both clusters, we found a high functional similarity between the same species coming from different vineyards.

The abundance of *Hanseniaspora* reads was found to be higher in samples W3 (72.0%) and W9 (44.2%) in vineyards 4 and 5, respectively, which had alcohol levels of ca. 1%. At the next time point sampled (equivalent to around 4% alcoholic level), we observed a decrease in relative abundance: W4 (8.2%) and W10 (22.8%) (Figure 3, Supplementary Table 6). Furthermore, *Saccharomyces* drastically increased by 27 fold in vineyard 4, and 3 fold in vineyard 5, between the 1% to 4% alcoholic level (Supplementary Table 6). This indicated the beginning of its early dominance in alcoholic fermentation. Of the other non-*Saccharomyces* yeasts solely found in vineyard 5 bins,

Metschnikowia had a similar trend as Hanseniaspora with the increase in alcoholic level (Supplementary Table 6).

Interestingly, it was observed that a higher total number of bacterial related contigs, as well as the percentage of read recruitment was obtained while a lower number of genes was identified in vineyard 4 than vineyard 5 (Supplementary Table 6). This might be due to a more diverse bacterial diversity in vineyard 4 than vineyard 5. Additionally, a decreasing trend in percentage of recruitment of all bacterial genomic bins was observed together with the increase in alcoholic level in vineyard 4. This followed the observation in the previous subsection with the increase of *Saccharomyces* activity. The *Metschnikowia* bin had 97 unique KOs, with 14 of those characterised as NADH dehydrogenase belonging to oxidative phosphorylation.

3.3 Alcoholic fermentation biodiversity

We explored the effect of alcohol level on microbial biodiversity using both the metabarcoding and metagenomics datasets (See Supplementary Tables 4 and 5 for details). Overall biodiversity was found to decrease during alcoholic fermentation. Specifically, the α-diversity, estimated using decomposed Hill numbers, clearly decreased with the fermentation progress (as measured in alcohol percentage) with both metabarcoding (Figure 2C) and metagenomic results (Figure 4A). Interestingly, the variance of Hill numbers generated and the α-diversity was found to increase with metagenomic filtered reads (Figure 4A, Supplementary Table 4), suggesting increased species diversity. This observation was further examined by focusing the analysis on a single vineyard (vineyard 4), using the shotgun metagenomic data generated at the four different alcoholic fermentation stages (W3-W6, Table 1). The assigned taxonomies were generated by mapping reads from these samples to a curated database (Supplementary Table 4b). In order to identify microbes that followed similar trajectories across the fermentation, affinity propagation clustering was applied on the relative abundances of species using Pearson correlation.

Five clusters were found (Figure 4B), and more in-depth investigation is shown in Supplementary Figure 3. The first contained species which showed an increase in their abundances during the fermentation (*S. cerevisiae*) (Supplementary Figure 3A). This principally implied that they were growing and active during fermentation. A cluster that included *Vitis vinifera*, *Botrytis cinerea* and *Erysiphe necator* also showed a slight decrease at the end of alcoholic fermentation (Supplementary Figure 3B). This indicated that they were relatively stable during the fermentation progress, while the relative amount of materials was inversely related to *Saccharomyces* growth phase during the process, which is quite understandable for grapevine DNA. A third, ambiguous, cluster showed a drastic decrease from the start, and no further increase until the end of fermentation (Supplementary Figure 3C). *Hanseniaspora* dominated in this cluster that mainly consisted of non-*Saccharomyces* yeasts that also acted similarly (Supplementary Figure 3C). This suggested that these species were active in the early fermentation but perish as the fermentation progresses. The fourth cluster showed a drastic decrease followed by an increase during the end of wine fermentation (Supplementary Figure 3D). This group was found to consist of both bacteria and fungi. The bacteria found in this group have been observed to relate to contamination originating from the commercial DNA extraction kit (Salter et al., 2014). The final (fifth) cluster with mainly *Micrococcus*, showed a decrease followed by an increase and ending up higher than in the start of fermentation (Figure 4C).

As the fifth cluster showed unexpected behaviour, further investigation was performed, through functional analysis on the ORFs generated from the metagenomic sequencing. Using the eggNOG taxonomic annotation, we separated each sample based on functions derived from eukaryotes, fungi, bacteria, virus and opisthokonts. The majority of the observed functions derived from fungi and bacteria (Figure 4D). A high number of functional counts that belonged to the bacteria were observed

at both the start and end of the fermentation; however, functional counts for fungi remained similar (Figure 4D). Therefore, the bacteria were further investigated by categorising these genes to different COGs. The eggNOG functional classes allowed investigation of gene functional categories of selected microbe groups (Figure 4E). A high proportion of the genes could not be resolved and were assigned with unknown function. The read recruitment on five genomic bins yielded from assembly and binning were also investigated. The *Hanseniaspora* and *Saccharomyces* bins performed as expected during fermentation. When looking into the relative abundance of recruited reads, a bin assigned to unknown Actinobacteria was observed to be increased in abundance in the last sampling date of the fermentation. This bin had 1969 genes identified with a 82.01% completeness and 0.72% redundancy (1.87 total size (Mb) and 319 contigs) (Supplementary Table 6).

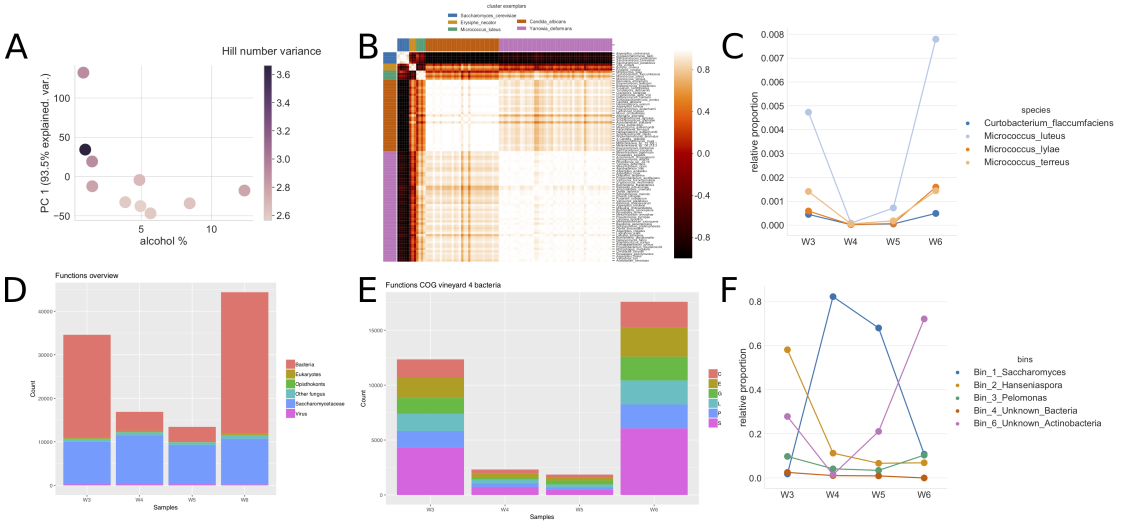


Figure 4. (A) Alcohol percentage and relationship to α-diversity of metagenomic sequenced samples of mapped reads show a decrease in $\boldsymbol{\alpha}$-diversity while gaining again in the end of fermentation. The α-diversity is shown as a first principal component of decomposed Hill numbers with q-values between 0 and 3. Color indicates change in variances of log10 fold Hill numbers. **(B)** A heatmap using pairwise clustering of mapped reads assigned to species. The assigned taxonomies generated from mapping to raw reads of the four samples from vineyard 4 to a curated database were clustered to observe patterns using affinity propagation of Pearson correlation similarities of the mapped reads assigned to species. Five clusters were found with corresponding cluster exemplars (blue: *Saccharomyces cerevisiae*, orange: *Erysiphe necator*, green: *Micrococcus luteus*, red: *Candida albicans*, purple: *Yarrowia deformans*). **(C)** Lineplots of the fifth clustered group of *Micrococcus* in **(4B)** show that these mapped reads assigned to species decrease during alcoholic fermentation, but also start increasing again in the end. **(D)** An overview of functional analysis of different groups: bacteria, eukaryotes, opisthokonts, other fungi, Saccharomycetaceae and viruses based on the read counts in vineyard 4 during alcoholic fermentation. Bacterial gene counts are observed to be most affected by the stage of fermentation, while other groups remain stable. **(E)** An overview of the changes in the bacterial functions to orthologous groups (COG) in vineyard 4. The COG groups are C: Energy production and conversion, E: Amino acid transport and metabolism, G: Carbohydrate metabolism and transport, L: Replication, recombination and repair, P: Inorganic ion transport and metabolism, and S: Function unknown. **(F)** The relative abundance of binned read recruitment of Anvi'o results. The size of dot increases with the increase in percentage of recruitment. The *Saccharomyces* bin is observed to grow during fermentation, while the unknown Actinobacteria bin is observed to increase in the end of fermentation.

3.4 Alcoholic fermentation biodiversity

In order to compare the level of fungal cells between samples, we performed qPCR targeting the ITS2 gene. Our data indicate that the overall level of fungal cells was lower in stuck samples (Supplementary Figure 4). We subsequently used the metabarcoding data to explore the microbial community differences between the stuck versus normally behaving ferments. Differential abundance analysis of the data enabled us to identify six OTUs that showed significant differences between the two phenotypes. In particular, two OTUs (a *Starmerella* species and an unknown fungal OTU) were significantly more abundant in the stuck samples (Figure 2B). The normal ferments contained a significantly higher abundance of four Ascomycete OTUs, of which three subsequently were classified to Saccharomycetales order (*Saccharomyces*, *Metschnikowia*, unknown Saccharomycetales OTU), while the fourth one remained unresolved. For stuck samples, the relative abundance of *Starmerella* species was observed to increase in each sampling time relating to longitudinal direction (Figure 1), whereas, *Saccharomyces* tended to dominate the normal fermentation behaviour phenotype. Chemical analyses identified several interesting observations on the stuck samples. Firstly, the fermentation speed was slower, and all fructose was consumed before glucose. Secondly, the glycerol concentration when less than 20 g/L of fructose remained was observed to be higher than in the normal performing samples. This suggested that when the fungal cells in the community was lower, the microbial community in the stuck samples preferred fructose over glucose.

4 Discussion

Our application of metabarcoding and shotgun metagenomic sequencing techniques to spontaneous Riesling ferments from four German vineyards enabled us to shed new insights into a number of questions of relevance to winemaking. We discuss each in further detail below.

4.1 Binning based approaches provide detailed insights into microbial differences between vineyards

Both metabarcoding and metagenomic approaches found that *Metschnikowia* drove the difference between vineyards. While the two methods offer similar information, differences in relative abundances might exist. For instance, one difference was suggested in the first shotgun sequencing paper on wine, where a *Metschnikowia* abundance bias was found between shotgun analysis and ITS2 marker gene, with ITS2 marker gene overestimating (Sternes et al., 2017). Further research and comparisons of relevant methodologies and workflows will be needed.

One of the main shortcomings of HTS-based approaches relates to limitations with current reference databases. It has been long acknowledged that ITS and other fungal marker gene regions have limited potential for resolving species identity (Scorzetti et al., 2002; Stefanini and Cavalieri, 2018), and metagenomic reference databases face similar challenges when assigning taxonomic labels to metagenomic DNA sequences. For example, we found that several *Saccharomyces* species besides *S. cerevisiae* were mapped in our data, namely S. *paradoxus* and *S. pastorianus* (Supplementary Table 4b). Since previous studies have reported that *S. paradoxus* is rarely found in wine fermentations (Knight et al., 2018b; Sicard and Legras, 2011), and the close evolutionary relationship between *S. cerevisiae* and *S. paradoxus* (Borneman and Pretorius, 2015), it is possible that our mapping to *S. paradoxus* is an artefact driven by close sequence homology (Naumova et al., 2005).

Additionally, most of the metagenomic algorithms and annotations are built mainly for prokaryotes and some included archaea instead eukaryotic taxa (West et al., 2018). We see this particularly with the assembly of overlapping reads into continuous or semi-continuous genome fragments and the results from the completeness of binning. We also note that taxonomic assignments to genus level are not particularly satisfying with regards to obtaining useful insights into the fermentation process, thus, resolution to species or even strain level is needed. Studies on bacterial communities have

shown this is possible in theory, although challenging (Papudeshi et al., 2017; Segata, 2018). For eukaryotes, more work is needed especially as multiple strains or species maybe be included in a contig due to current challenges in distinguishing between related community members in both the assembly and binning processes (Imelfort et al., 2014; Luo et al., 2015; Parks et al., 2017).

Although using metagenomic approaches we were also able to provide functional insights, this also faces a major challenge. This is due to the relatively poor state of annotation for most bacterial and fungal genomes. Thus, while the potential benefits of (meta)genomic analysis are clear, inference is based on a rather shaky foundation. We observe this with the *Metschnikowia* bin, which has multiple unique functions for NADH dehydrogenases which are mapped to human diseases in KEGG, due to its oxidative phosphorylation (Schägger and Ohm, 1995). These functions also relate to the iron metabolism/electron transport/respiration as it has been suggested that these functions are up-regulated by MarR-like protein PchR in pulcherriminic acid biosynthetic pathway of *Bacillus subtilis*. (Randazzo et al., 2016). This pulcherriminic acid present in the medium while depleting the iron is suggested to be the main reason for the role of *Metschnikowia* as a biocontrol agent against other non-*Saccharomyces* in wine (Oro et al., 2014), This could therefore relate to bacteria present in wine, as a higher number of bacterial genes but a lower total number of bacterial related contigs, was identified in vineyard 5 compared to vineyard 4 (Supplementary Table 6). This could help to create a less diverse environment for *Saccharomyces* to dominate and complete the alcoholic fermentation easier (Boynton and Greig, 2016).

4.2 Alcoholic fermentation biodiversity generally decreases but novel bacteria increase is found to increase using metagenomics

A decline in biodiversity during alcoholic fermentation was observed in both amplicon and mapping raw reads from metagenomic analysis. Previously with metabarcoding, it has been found that some bacteria become more abundant in the later stages of fermentations (Bokulich et al., 2015).

In this study, we noticed an increase in unidentified Actinobacteria, both when recruiting the read results to binned data, and in the functional assignments, on vineyard 4. The raw reads mapping to the curated database suggested this bin might be related to *Micrococcus* genus. Actinobacteria in general has been observed in wine and in Riesling (Piao et al., 2015) previously but without further interpretation (Bokulich et al., 2013; Marzano et al., 2016; Portillo and Mas, 2016). *Micrococcus*, in particular, has been isolated previously from beer (Priest, 1999), cheese (Fontana et al., 2010) and other fermented foods where *Micrococcus* is known to both produce and degrade biogenic amines through amino acid decarboxylases (Leuschner et al., 1998; Voigt and Eitenmiller, 1977) as well as produce volatile sulphur compounds (Bonnarme et al., 2000).

The dynamic changes of the community composition during the fermentation, could affect the detection threshold of the bacteria, which plays a key role in the unique functional counts. It is widely known that *Saccharomyces* dominates the microbial community during alcoholic fermentation, and that the total microbial biodiversity decreases. These two patterns could be driven by two possible explanations. Firstly, due to the proportional increase of *Saccharomyces* to other microbes, given that we maintained a relatively constant sequencing depth across the samples, we may well be missing less abundant microbes (such as bacteria). As a result, a big decline in abundance and identified functions of bacteria would be observed when *Saccharomyces* levels are at their peak (Figure 4F). A second explanation could be that the observed decline in bacteria is the actual biological behavior, for example if bacteria enter a survival mode with drastically decreased activity alongside the rapid growth of *Saccharomyces*. Towards the end of alcoholic fermentation, the activity of *Saccharomyces* slows down and provides space for the growth of bacteria. This possibility is supported by the observed functions of other groups, like eukaryotes, virus and opisthokonts, where no clear decline is observed but rather a steady identification rate of their functions across the fermentation (Figure 4D).

Thus, we hypothesise that the functionality and the increase in actual species is a real observation, although further validation with multi-omic studies is needed.

4.3 Starmerella and stuck fermentation behaviour

Our analyses demonstrated that stuck fermentation behaviour during alcoholic fermentation was associated with a presence of fructophilic *Starmerella* and absence of *Saccharomyces*. These findings are consistent with those from previous studies in which *Starmerella* (synonym *Candida zemplinina*) had been investigated because of its' known fructophilic characters, aroma profile as well as lower ethanol production, and elevated glycerol contents (Ciani and Comitini, 2015; Giaramida et al., 2013; Masneuf-Pomarede et al., 2015). That it has a reduced rate of fermentation (Aponte and Blaiotta, 2016) and thus requires need for *S. cerevisiae* to finish the alcoholic fermentation (Masneuf-Pomarede et al., 2015) has also been suggested based inoculation studies. In addition to providing the first HTS based insights into this, our results provide the first evidence of these characteristics in samples taken from natural winemaking environments. Our dataset also enabled us to observe that samples with more *Starmerella* had reduced levels of *Botrytis*. This is intriguing given that *Starmerella* was isolated the first time from sweet botrytised wines with high fructose-glucose ratio (Csoma and Sipiczki, 2008; Sipiczki, 2004). Thus we suggest that further investigation into the relationship between the two using HTS techniques may be of interest, as specific links between *Starmerella* and winemaking have not currently been established (Masneuf-Pomarede et al., 2015).

While HTS tools offer unprecedented potential, we see lot of challenges and possibilities when combining HTS tools together with other 'omics methods such as metabolomics, or metatranscriptomics, as recently suggested (Sirén et al., 2019). However, the value of multi-omics is the ability to examine a large number of samples that is needed to draw robust statistically valid conclusions. Thus, classic issues that arise during experimental design (Knight et al., 2018a), such as the challenges in metadata and sampling size, and issues of mechanism and causation (Fischbach, 2018) need to be kept in mind.

5 Conclusion

In summary, we demonstrate the power of HTS-based tools in characterising microbial community differences and fermentation population dynamics. While the metagenomics and metabarcoding approaches were found to deliver similar results, the former provides a more in-depth understanding given it offers an untargeted taxonomical analysis, as well as enabling insights at the functional level. Ultimately as more such studies appear, we anticipate that they will catalyse significant further future wine microbial research.

6 Conflict of Interest

Author JHS was employed by company Chr. Hansen. All other authors declare no competing interests.

7 Author Contributions

KS and SSTM conceived the ideas, designed the experiment and co-wrote the manuscript with inputs from all authors. KS performed the sample preparation and fermentation trial set up and data collection. SSTM carried out the laboratory work and data collection. KS and SSTM performed the analyses and interpretations of metabarcoding data. CM led the analysis and interpretation of metagenomic data with the inputs from KS and SSTM. KS and CM curated the wine metagenomic database. CC developed the shotgun library construction protocols and performed the protocol comparison with SSTM. JHS, DM, UF and MTPG oversaw the study and edited the manuscript. All authors read and approved the final version of the manuscript.

8 Funding

The research was funded by the Horizon 2020 Programme of the European Commission within the Marie Skłodowska-Curie Innovative Training Network "MicroWine" (grant number 643063), as well as the Innovation Fund Denmark project "FoodTranscriptomics" (grant number 6150-00033B).

9 Acknowledgments

The authors would like to thank Engela Stadler in assisting the fermentation trial set up, Pascal Wegmann-Herr for helpful discussions for experimental design, Martin René Ellegaard in assisting DNA extraction, Mikkel Sinding for providing the wolf samples in metagenomic library construction methods comparison, Anna Fotakis for advice in data analysis, the wine team at Chr. Hansen A/S and the collaborating wine estates in Pfalz.

10 References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106. doi:10.1186/gb-2010-11-10-r106.

Andrews, S. (2010). FastQC A quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed December 10, 2018].

Aponte, M., and Blaiotta, G. (2016). Potential role of yeast strains isolated from grapes in the production of Taurasi DOCG. Front. Microbiol. 7, 809. doi:10.3389/fmicb.2016.00809.

Bagheri, B., Bauer, F. F., and Setati, M. E. (2017). The Impact of Saccharomyces cerevisiae on a wine yeast consortium in natural and inoculated fermentations. Front. Microbiol. 8, 1988. doi:10.3389/fmicb.2017.01988.

Belda, I., Zarraonaindia, I., Perisin, M., Palacios, A., and Acedo, A. (2017). From vineyard soil to wine fermentation: microbiome approximations to explain the ``terroir'' concept. Front. Microbiol. 8, 821.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. 57, 289–300. Available at: http://www.jstor.org/stable/2346101.

Bisson, L. F., Joseph, C. M. L., and Domizio, P. (2017). "Yeasts," in Biology of Microorganisms on Grapes, in Must and in Wine, eds. H. König, G. Unden, and J. Fröhlich (Cham: Springer International Publishing), 65–101.

Bodenhofer, U., Kothmeier, A., and Hochreiter, S. (2011). APCluster: an R package for affinity propagation clustering. Bioinformatics 27, 2463–2464. doi:10.1093/bioinformatics/btr406.

Bokulich, N. A., Collins, T. S., Masarweh, C., Allen, G., Heymann, H., Ebeler, S. E., et al. (2016). Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. MBio 7.

Bokulich, N. A., Ohta, M., Richardson, P. M., and Mills, D. A. (2013). Monitoring seasonal changes in winery-resident microbiota. PLoS One 8, e66437. doi:10.1371/journal.pone.0066437.

Bokulich, N. A., Swadener, M., Sakamoto, K., Mills, D. A., and Bisson, L. F. (2015). Sulfur dioxide treatment alters wine microbial diversity and fermentation progression in a dose-dependent fashion. Am. J. Enol. Vitic. 66, 73–79. doi:10.5344/ajev.2014.14096.

Bokulich, N. A., Thorngate, J. H., Richardson, P. M., and Mills, D. A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. Proc. Natl. Acad. Sci. U. S. A. 111, E139–48. doi:10.1073/pnas.1317377110.

Bonnarme, P., Psoni, L., and Spinnler, H. E. (2000). Diversity of L-methionine catabolism pathways in cheese-ripening bacteria. Appl. Environ. Microbiol. 66, 5514–5517. Available at: https://www.ncbi.nlm.nih.gov/pubmed/11097940.

Borneman, A. R., and Pretorius, I. S. (2015). Genomic insights into the Saccharomyces sensu stricto complex. Genetics 199, 281–291. doi:10.1534/genetics.114.173633.

Boynton, P. J., and Greig, D. (2016). Species richness influences wine ecosystem function through a dominant species. Fungal Ecol. 22, 61–72. doi:10.1016/j.funeco.2016.04.008.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. doi:10.1038/nmeth.f.303.

Ciani, M., Beco, L., and Comitini, F. (2006). Fermentation behaviour and metabolic interactions of multistarter wine yeast fermentations. Int. J. Food Microbiol. 108, 239–245. doi:10.1016/j.ijfoodmicro.2005.11.012.

Ciani, M., and Comitini, F. (2015). Yeast interactions in multi-starter wine fermentation. Current Opinion in Food Science 1, 1–6. doi:10.1016/j.cofs.2014.07.001.

Cokelaer, T., Pultz, D., Harder, L. M., Serra-Musach, J., and Saez-Rodriguez, J. (2013). BioServices: a common Python package to access biological Web Services programmatically. Bioinformatics 29, 3241–3242. doi:10.1093/bioinformatics/btt547.

Csoma, H., and Sipiczki, M. (2008). Taxonomic reclassification of Candida stellata strains reveals frequent occurrence of Candida zemplinina in wine fermentation. FEMS Yeast Res. 8, 328–336. doi:10.1111/j.1567-1364.2007.00339.x.

Delmont, T. O., Eren, A. M., Maccario, L., Prestat, E., Esen, Ö. C., Pelletier, E., et al. (2015). Reconstructing rare soil microbial genomes using in situ enrichments and metagenomics. Front. Microbiol. 6, 358. doi:10.3389/fmicb.2015.00358.

Díaz, C., Badalyan, G., and Bücking, M. (2018). "Molecular techniques for the detection and identification of yeasts in wine," in Molecular Techniques in Food Biology, eds. A. F. El Sheikha, R. Levin, and J. Xu (Chichester, UK: John Wiley & Sons, Ltd), 323–340. doi:10.1002/9781119374633.ch14.

Di Maro, E., Ercolini, D., and Coppola, S. (2007). Yeast dynamics during spontaneous wine fermentation of the Catalanesca grape. Int. J. Food Microbiol. 117, 201–210. doi:10.1016/j.ijfoodmicro.2007.04.007.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. doi:10.1093/bioinformatics/btq461.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998. doi:10.1038/nmeth.2604.

Fang, C., Zhong, H., Lin, Y., Chen, B., Han, M., Ren, H., et al. (2018). Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. Gigascience 7, 1–8.

Feld, L., Nielsen, T. K., Hansen, L. H., Aamand, J., and Albers, C. N. (2016). Establishment of bacterial herbicide degraders in a rapid sand filter for bioremediation of phenoxypropionate-polluted groundwater. Appl. Environ. Microbiol. 82, 878–887. doi:10.1128/AEM.02600-15.

Fischbach, M. A. (2018). Microbiome: Focus on Causation and Mechanism. Cell 174, 785–790. doi:10.1016/j.cell.2018.07.038.

Fleet, G. H. (1990). Growth of yeasts during wine fermentations. J. Wine Res. 1, 211–223. doi:10.1080/09571269008717877.

Fleet, G. H. (2003). Yeast interactions and wine flavour. Int. J. Food Microbiol. 86, 11–22. doi:10.1016/S0168-1605(03)00245-9.

Fontana, C., Cappa, F., Rebecchi, A., and Cocconcelli, P. S. (2010). Surface microbiota analysis of Taleggio, Gorgonzola, Casera, Scimudin and Formaggio di Fossa Italian cheeses. Int. J. Food Microbiol. 138, 205–211. doi:10.1016/j.ijfoodmicro.2010.01.017.

Frey, B. J., and Dueck, D. (2007). Clustering by passing messages between data points. Science 315, 972–976. doi:10.1126/science.1136800.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., et al. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat. Commun. 8, 1188. doi:10.1038/s41467-017-01312-x.

Ganucci, D., Guerrini, S., Mangani, S., Vincenzini, M., and Granchi, L. (2018). Quantifying the Effects of ethanol and temperature on the fitness advantage of predominant Saccharomyces cerevisiae strains occurring in spontaneous wine fermentations. Front. Microbiol. 9, 1563.

Giaramida, P., Ponticello, G., Di Maio, S., Squadrito, M., Genna, G., Barone, E., et al. (2013). Candida zemplinina for production of wines with less alcohol and more glycerol. S. Afr. J. Enol. Vitic. 34, 204–211. Available at: http://www.journals.ac.za/index.php/sajev/article/download/1095/404.

Grangeteau, C., Gerhards, D., Rousseaux, S., von Wallbrunn, C., Alexandre, H., and Guilloux-Benatier, M. (2015). Diversity of yeast strains of the genus Hanseniaspora in the winery environment: What is their involvement in grape must fermentation? Food Microbiol. 50, 70–77. doi:10.1016/j.fm.2015.03.009.

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. Ecology 54, 427–432. doi:10.2307/1934352.

HMMER (2015). HMMER: biosequence analysis using profile hidden Markov models. Available at: http://hmmer.org. [Accessed November 23, 2018].

Hofmann, H., Wickham, H., and Kafadar, K. (2017). Letter-value plots: boxplots for large data. J. Comput. Graph. Stat. 26, 469–477. doi:10.1080/10618600.2017.1305277.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. 9, 90–95. doi:10.1109/MCSE.2007.55.

Hyatt, D., LoCascio, P. F., Hauser, L. J., and Uberbacher, E. C. (2012). Gene and translation initiation site prediction in metagenomic sequences. Bioinformatics 28, 2223–2230. doi:10.1093/bioinformatics/bts429.

Ihrmark, K., Bödeker, I. T. M., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al. (2012). New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol. Ecol. 82, 666–677. doi:10.1111/j.1574-6941.2012.01437.x.

Imelfort, M., Parks, D., Woodcroft, B. J., Dennis, P., Hugenholtz, P., and Tyson, G. W. (2014). GroopM: an automated tool for the recovery of population genomes from related metagenomes. PeerJ 2, e603.

Jolly, N. P., Varela, C., and Pretorius, I. S. (2014). Not your ordinary yeast: non-Saccharomyces yeasts in wine production uncovered. FEMS Yeast Res. 14, 215–237.

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018a). Best practices for analysing microbiomes. Nat. Rev. Microbiol. 16, 410–422.

Knight, S. J., Klaere, S., Morrison-Whittle, P., and Goddard, M. R. (2018b). Fungal diversity during fermentation correlates with thiol concentration in wine. Aust. J. Grape Wine Res. 24, 105–112. doi:10.1111/ajgw.12304.

Krueger, F. (2012). Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. Available at: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed December 10, 2018].

Leuschner, R. G., Heidel, M., and Hammes, W. P. (1998). Histamine and tyramine degradation by food fermenting microorganisms. Int. J. Food Microbiol. 39, 1–10. doi:10.1016/S0168-1605(97)00109-8.

Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., et al. (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11. doi:10.1016/j.ymeth.2016.02.020.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN]. Available at: http://arxiv.org/abs/1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

Liu, P.-T., Lu, L., Duan, C.-Q., and Yan, G.-L. (2016). The contribution of indigenous non-Saccharomyces wine yeast to improved aromatic quality of Cabernet Sauvignon wines by spontaneous fermentation. LWT - Food Science and Technology 71, 356–363. doi:10.1016/j.lwt.2016.04.031.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. doi:10.1186/s13059-014-0550-8.

Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci. 3, e104. doi:10.7717/peerj-cs.104.

Luo, C., Knight, R., Siljander, H., Knip, M., Xavier, R. J., and Gevers, D. (2015). ConStrains identifies microbial strains in metagenomic datasets. Nat. Biotechnol. 33, 1045–1052. doi:10.1038/nbt.3319.

Martini, A. (1993). Origin and domestication of the wine yeast Saccharomyces cerevisiae. J. Wine Res. 4, 165–176. doi:10.1080/09571269308717966.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12. doi:10.14806/ej.17.1.200.

Marzano, M., Fosso, B., Manzari, C., Grieco, F., Intranuovo, M., Cozzi, G., et al. (2016). Complexity and dynamics of the winemaking bacterial communities in berries, musts, and wines from Apulian grape cultivars through time and space. PLoS One 11, e0157383. doi:10.1371/journal.pone.0157383.

Masneuf-Pomarede, I., Juquin, E., Miot-Sertier, C., Renault, P., Laizet, Y. 'han, Salin, F., et al. (2015). The yeast Starmerella bacillaris (synonym Candida zemplinina) shows high genetic diversity in winemaking environments. FEMS Yeast Res. 15, fov045. doi:10.1093/femsyr/fov045.

McKinney, W. (2011). pandas: a foundational Python library for data analysis and statistics. PyHPC 2011, 1–9.

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8, e61217. doi:10.1371/journal.pone.0061217.

Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 11257. doi:10.1038/ncomms11257.

Morgan, S. C., Tantikachornkiat, M., Scholl, C. M., Benson, N. L., Cliff, M. A., and Durall, D. M. (2018). The effect of sulfur dioxide addition at crush on the fungal and bacterial communities and the sensory attributes of Pinot gris wines. Int. J. Food Microbiol. 290, 1–14. doi:10.1016/j.ijfoodmicro.2018.09.020.

Morrison-Whittle, P., and Goddard, M. R. (2018). From vineyard to winery: a source map of microbial diversity driving wine fermentation. Environ. Microbiol. 20, 75–84.

Murat Eren, A., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3, e1319. doi:10.7717/peerj.1319.

Naumova, E. S., Naumov, G. I., Masneuf-Pomarède, I., Aigle, M., and Dubourdieu, D. (2005). Molecular genetic study of introgression between Saccharomyces bayanus and S. cerevisiae. Yeast 22, 1099–1115. doi:10.1002/yea.1298.

Ocón, E., Garijo, P., Sanz, S., Olarte, C., López, R., Santamaría, P., et al. (2013). Screening of yeast mycoflora in winery air samples and their risk of wine contamination. Food Control 34, 261–267. doi:10.1016/j.foodcont.2013.04.044.

Oliver, A. K., Brown, S. P., Callaham, M. A., and Jumpponen, A. (2015). Polymerase matters: non-proofreading enzymes inflate fungal community richness estimates by up to 15 %. Fungal Ecol. 15, 86–89. doi:10.1016/j.funeco.2015.03.003.

Oro, L., Ciani, M., and Comitini, F. (2014). Antimicrobial activity of Metschnikowia pulcherrima on wine yeasts. J. Appl. Microbiol. 116, 1209–1217. doi:10.1111/jam.12446.

Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2017). Biostrings: Efficient manipulation of biological strings. R Package Version 2.

Papudeshi, B., Haggerty, J. M., Doane, M., Morris, M. M., Walsh, K., Beattie, D. T., et al. (2017). Optimizing and evaluating the reconstruction of Metagenome-assembled microbial genomes. BMC Genomics 18, 915. doi:10.1186/s12864-017-4294-1.

Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol 2, 1533–1542.

Pasteur, L. (1872). Nouvelles expériences pour démontrer que le germe de la levure qui fait le vin provient de l'extérieur des grains de raisin. CR Searces Acad. Sci. Paris 75, 781–793.

Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28, 1420–1428. doi:10.1093/bioinformatics/bts174.

Pérez-Martín, F., Seseña, S., Fernández-González, M., Arévalo, M., and Palop, M. L. (2014). Microbial communities in air and wine of a winery at two consecutive vintages. Int. J. Food Microbiol. 190, 44–53. doi:10.1016/j.ijfoodmicro.2014.08.020.

Piao, H., Hawley, E., Kopf, S., DeScenzo, R., Sealock, S., Henick-Kling, T., et al. (2015). Insights into the bacterial community and its temporal succession during the fermentation of wine grapes. Front. Microbiol. 6, 809. doi:10.3389/fmicb.2015.00809.

Portillo, M. del C., and Mas, A. (2016). Analysis of microbial diversity and dynamics during wine fermentation of Grenache grape variety by high-throughput barcoding sequencing. LWT - Food Science and Technology 72, 317–321. doi:10.1016/j.lwt.2016.05.009.

Powell, S., Forslund, K., Szklarczyk, D., Trachana, K., Roth, A., Huerta-Cepas, J., et al. (2014). eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res. 42, D231–9. doi:10.1093/nar/gkt1253.

Pretorius, I. S. (2000). Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. Yeast 16, 675–729. doi:3.0.CO;2-B">10.1002/1097-0061(20000615)16:8<675::AID-YEA585>3.0.CO;2-B.

Priest, F. G. (1999). "Gram-positive brewery bacteria," in Brewing Microbiology, eds. F. G. Priest and I. Campbell (Boston, MA: Springer US), 127–161. doi:10.1007/978-1-4684-0038-0_5.

Randazzo, P., Aubert-Frambourg, A., Guillot, A., and Auger, S. (2016). The MarR-like protein PchR (YvmB) regulates expression of genes involved in pulcherriminic acid biosynthesis and in the initiation of sporulation in Bacillus subtilis. BMC Microbiol. 16, 190. doi:10.1186/s12866-016-0807-3.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. doi:10.7717/peerj.2584.

Romano, P., Fiore, C., Paraggio, M., Caruso, M., and Capece, A. (2003). Function of yeast species and strains in wine flavour. Int. J. Food Microbiol. 86, 169–180. doi:10.1016/S0168-1605(03)00290-3.

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87. doi:10.1186/s12915-014-0087-z.

Sarkar, D. (2008). Lattice: multivariate data visualization with R. Springer Science & Business Media Available at: https://market.android.com/details?id=book-gXxKFWkE9h0C.

Schägger, H., and Ohm, T. G. (1995). Human diseases with defects in oxidative phosphorylation: 2. F1F0 ATP-synthase defects in Alzheimer disease revealed by blue native polyacrylamide gel electrophoresis. Eur. J. Biochem. 227, 916–921. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1432-1033.1995.0916p.x.

Scorzetti, G., Fell, J. W., Fonseca, A., and Statzell-Tallman, A. (2002). Systematics of basidiomycetous yeasts: a comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. FEMS Yeast Res. 2, 495–517. Available at: https://www.ncbi.nlm.nih.gov/pubmed/12702266.

Segata, N. (2018). On the road to strain-resolved comparative metagenomics. mSystems 3. doi:10.1128/mSystems.00190-17.

Sicard, D., and Legras, J.-L. (2011). Bread, beer and wine: yeast domestication in the Saccharomyces sensu stricto complex. C. R. Biol. 334, 229–236. doi:10.1016/j.crvi.2010.12.016.

Sipiczki, M. (2004). Species identification and comparative molecular and physiological analysis of Candida zemplinina and Candida stellata. J. Basic Microbiol. 44, 471–479. doi:10.1002/jobm.200410449.

Sirén, K., Mak, S. S. T., Fischer, U., Hansen, L. H., and Gilbert, M. T. P. (2019). Multi-omics and potential applications in wine production. Curr. Opin. Biotechnol. 56, 172–178. doi:10.1016/j.copbio.2018.11.014.

Sokal, R. R., and Michener, C. D. (1958). "A statistical method for evaluating systematic relationships," in University of Kansas science bulletin (University of Kansas), 1409–38. Available at: https://market.android.com/details?id=book-o1BlHAAACAAJ.

Stefanini, I., Albanese, D., Cavazza, A., Franciosi, E., De Filippo, C., Donati, C., et al. (2016). Dynamic changes in microbiota and mycobiota during spontaneous "Vino Santo Trentino" fermentation. Microb. Biotechnol. 9, 195–208. doi:10.1111/1751-7915.12337.

Stefanini, I., and Cavalieri, D. (2018). Metagenomic approaches to investigate the contribution of the vineyard environment to the quality of wine fermentation: potentials and difficulties. Front. Microbiol. 9, 991.

Sternes, P. R., Lee, D., Kutyna, D. R., and Borneman, A. R. (2017). A combined meta-barcoding and shotgun metagenomic analysis of spontaneous wine fermentation. Gigascience 6, 1–10.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., et al. (2003). The COG database: an updated version includes eukaryotes. BMC Bioinformatics 4, 41. doi:10.1186/1471-2105-4-41.

Tempère, S., Marchal, A., Barbe, J.-C., Bely, M., Masneuf-Pomarede, I., Marullo, P., et al. (2018). The complexity of wine: clarifying the role of microorganisms. Appl. Microbiol. Biotechnol. 102, 3995–4007.

Tenenbaum, D. (2018). KEGGREST: Client-side REST access to KEGG. R package version 1.23.0. Available at: https://bioconductor.org/packages/release/bioc/html/KEGGREST.html.

Torija, M. J., Rozès, N., Poblet, M., Guillamón, J. M., and Mas, A. (2001). Yeast population dynamics in spontaneous fermentations: comparison between two different wine-producing areas over a period of three years. Antonie Van Leeuwenhoek 79, 345–352. Available at: https://www.ncbi.nlm.nih.gov/pubmed/11816978.

Varela, C., and Borneman, A. R. (2017). Yeasts found in vineyards and wineries. Yeast 34, 111–128.

Voigt, M. N., and Eitenmiller, R. R. (1977). Production of tyrosine and histidine decarboxylase by dairy-related bacteria. J. Food Prot. 40, 241–245. doi:10.4315/0362-028X-40.4.241.

Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lukauskas, S., Gemperline, D. C., et al. (2017). mwaskom/seaborn: v0.8.1 (September 2017). doi:10.5281/zenodo.883859.

Werner, J. J., Zhou, D., Caporaso, J. G., Knight, R., and Angenent, L. T. (2012). Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J. 6, 1273–1276. doi:10.1038/ismej.2011.186.

West, P. T., Probst, A. J., Grigoriev, I. V., Thomas, B. C., and Banfield, J. F. (2018). Genome-reconstruction for eukaryotes from complex natural microbial communities. Genome Res. 28, 569–580. doi:10.1101/gr.228429.117.

White, T. J., Bruns, T., Lee, S., Taylor, J. L., and Others (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications 18, 315–322.

Wickham, H. (2007). Reshaping data with the reshape package. J. Stat. Softw. 21, 1–20. Available at: http://had.co.nz/reshape/introduction.pdf.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. Springer Available at: https://market.android.com/details?id=book-XgFkDAAAQBAJ.

Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 15, R46. doi:10.1186/gb-2014-15-3-r46.

Zepeda-Mendoza, M. L., Edwards, N. K., Madsen, M. G., Abel-Kistrup, M., Puetz, L., Sicheritz-Ponten, T., et al. (2018). Influence of Oenococcus oeni and Brettanomyces bruxellensis on wine microbial taxonomic and functional potential profile. Am. J. Enol. Vitic. 69, 321–333. doi:10.5344/ajev.2018.17092.

Zott, K., Thibon, C., Bely, M., Lonvaud-Funel, A., Dubourdieu, D., and Masneuf-Pomarede, I. (2011). The grape must non-Saccharomyces microbial community: impact on volatile thiol release. Int. J. Food Microbiol. 151, 210–215. doi:10.1016/j.ijfoodmicro.2011.08.026.

11 Supplementary Material

Supplementary file 1 - shotgun library construction method comparison

Supplementary tables

Supplementary figures

12 Data Availability Statement

The sequencing data were deposited to European Nucleotide Archive under study number: PRJEB30801 and ERS3017411-ERS3017414, ERS3017423-26, ERS3017435-38, ERS3017447-50, ERS3017459-62, ERS3017471-74, ERS3017483-86, ERS3017495-98 in study number: PRJEB29796.

13 Tables

Table 1. Summary of samples analysed. Each sample name can be used to identify the vineyard name (first 2 characters), biological replicate (a or b), and the sampling date in 2015 (last 2 numbers).

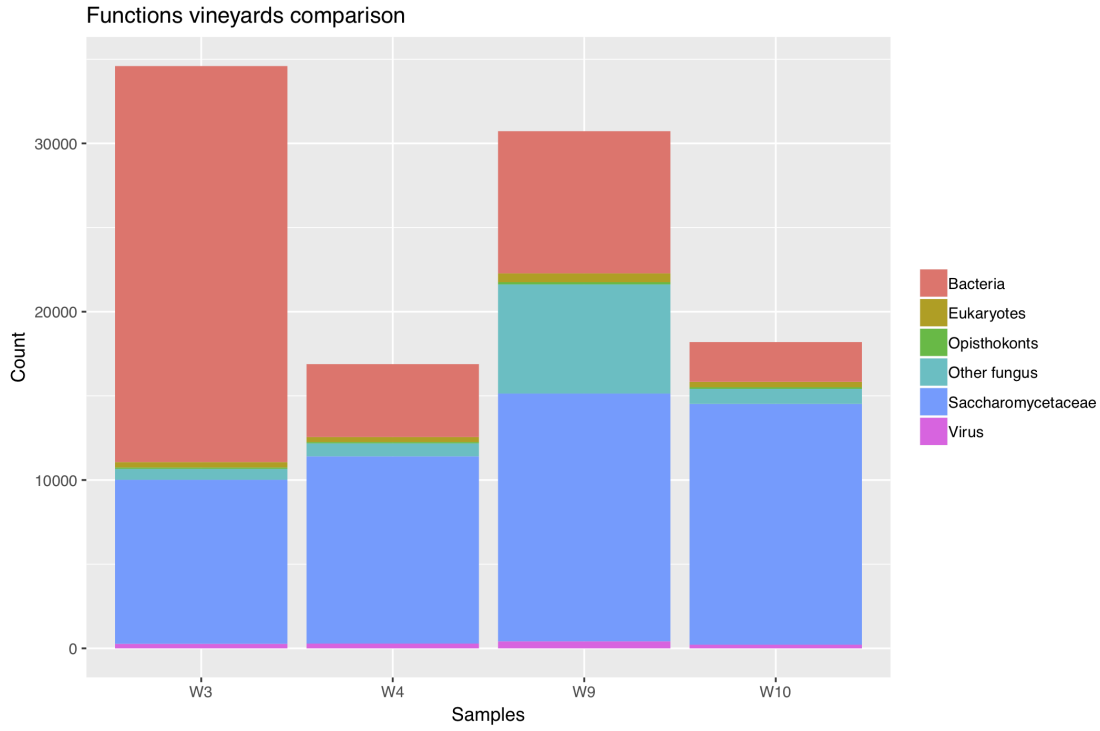
No.	Sample name	Metagenomic sample ID	Vineyard	Biological replicate	Sampling date	Alcoholic percentage
1	w2a109		2	a	09-Oct	0.5
2	w2b109		2	b	09-Oct	1.3
3	w3a109*	W1	3	a	09-Oct	1.5
4	w3b109		3	b	09-Oct	6
5	w4a109*	W3	4	a	09-Oct	1
6	w4b109		4	b	09-Oct	1
7	w5a109		5	a	09-Oct	0.22
8	w5b109		5	b	09-Oct	0.14
9	w2a112		2	a	12-Oct	1.11
10	w2b112		2	b	12-Oct	1.62
11	w3a112*	W2	3	a	12-Oct	5.65
12	w3b112		3	b	12-Oct	9.51
13	w4a112*	W4	4	a	12-Oct	4.93
14	w4b112		4	b	12-Oct	1.79
15	w5a112*	W7	5	a	12-Oct	0.87

16	w5b112*	W9	5	b	12-Oct	1.49
17	w2a114		2	a	14-Oct	1.13
18	w2b114		2	b	14-Oct	1.99
19	w3a114		3	a	14-Oct	8.18
20	w3b114		3	b	14-Oct	10.64
21	w4a114*	W5	4	a	14-Oct	8.44
22	w4b114		4	b	14-Oct	2.16
23	w5a114*	W8	5	a	14-Oct	3.85
24	w5b114*	W10	5	b	14-Oct	4.86
25	w2a120		2	a	20-Oct	1.65
26	w2b120		2	b	20-Oct	3.66
27	w3a120		3	a	20-Oct	11.41
28	w3b120		3	b	20-Oct	11.85
29	w4a120*	W6	4	a	20-Oct	12.34
30	w4b120		4	b	20-Oct	3.47
31	w5a120		5	a	20-Oct	10.5
32	w5b120		5	b	20-Oct	10.34

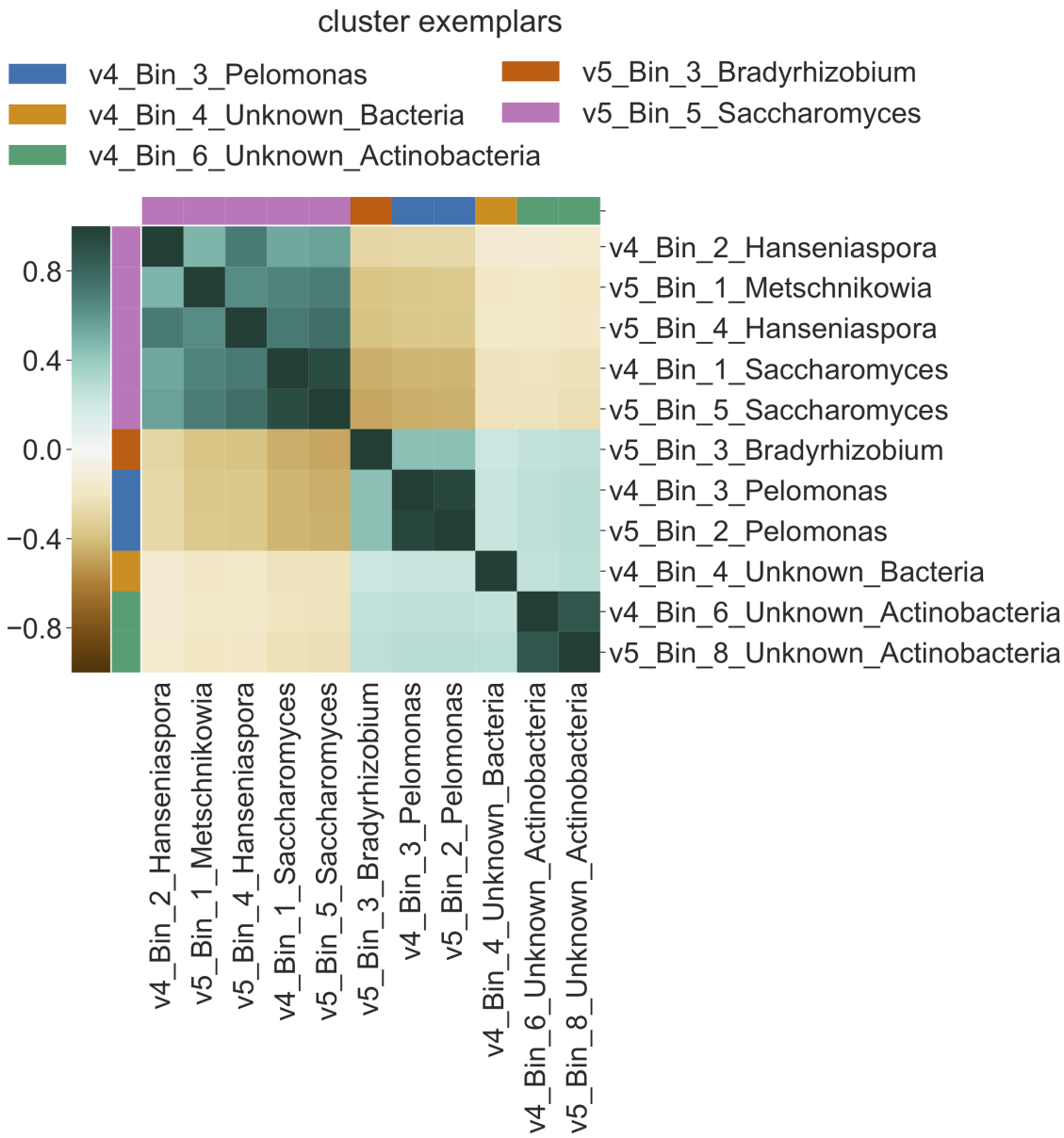
*Samples selected for metagenomic sequencing.

Supplementary information of Chapter 4

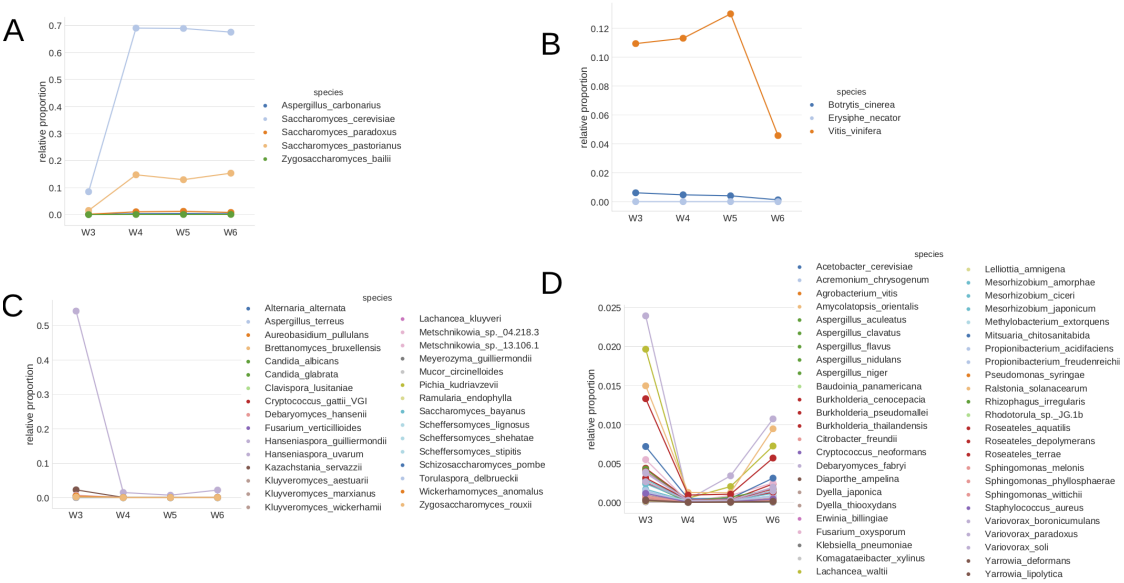
14 Supplementary figure captions



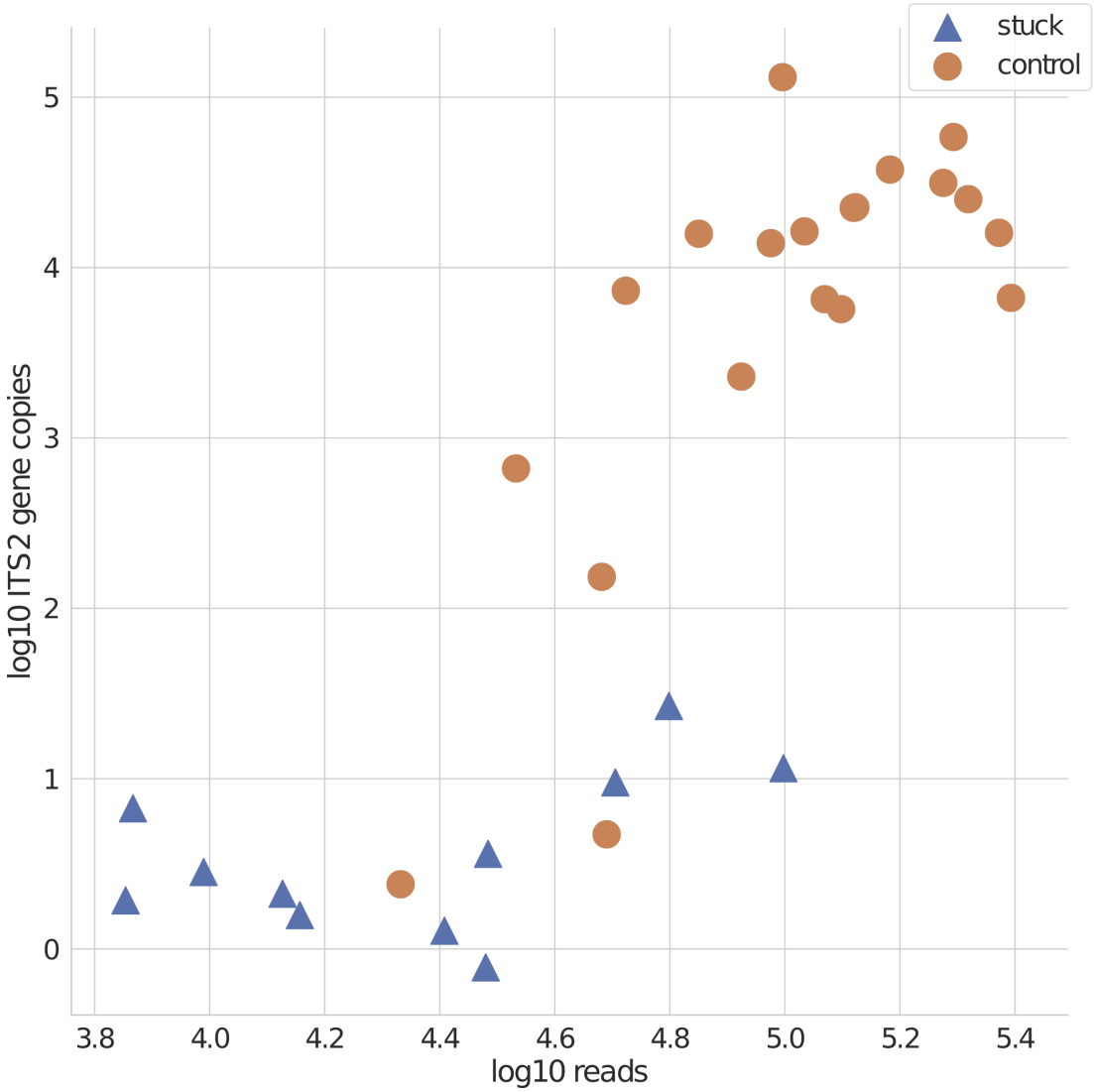
Supplementary Figure 1. Absolute counts of functions with taxonomy separation from eggNOG results with the top 6 groups included: bacteria, eukaryotes (without fungi), opisthokonts, other fungus, family Saccharomycetaceae and virus. Functions compared between vineyard 4 (W3 and W4) and vineyard 5 (W9 and W10) show abundances and differences caused mainly by two groups: bacteria and a collective group of "other fungus and Saccharomycetaceae".



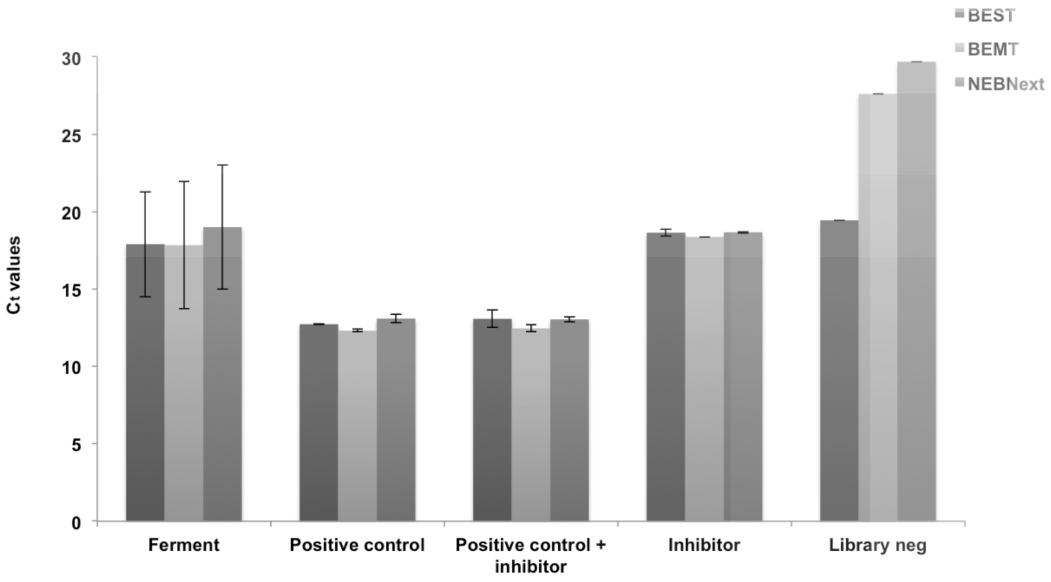
Supplementary Figure 2. Heatmap visualisation of the drafted genomic bins assigned with unique 6394 KEGG Orthology annotated functions using affinity propagation with Pearson correlation on the presence and absence table. Cluster colours are based on unsupervised clustering from affinity propagation.



Supplementary Figure 3 Affinity propagation clustering of mapped taxa using Pearson correlation. **(A)** The species which had an increase of their abundance during the fermentation were grouped to cluster 1. **(B)** Second cluster included *Vitis vinifera*, *Erysiphe necator* and *Botrytis cinerea*. **(C)** Third cluster was wine related yeasts driven cluster dominated with *Hanseniaspora uvarum*. The cluster shows a decrease from the start and no further increase until the end. **(D)** Fourth ambiguous cluster with yeast and bacteria.



Supplementary Figure 4. Metabarcoding performance. Sample read numbers from metabarcoding mapped against ITS2 gene copy numbers with samples colored and shaped according to fermentation behaviour. The data show that samples with stuck fermentation behaviour had the lowest amounts of both reads and ITS2 gene copy numbers. Blue triangle: stuck fermentation behaviour, orange circle: normal fermentation behaviour.



Supplementary Figure 5. Mean C_t values (± standard deviation) from qPCR of three extraction methods in different sample types. n=5 for ferment, while n=2 for other.

15 Supplementary table captions

Supplementary Table 1. Details of samples and measured wine parameters.

Supplementary Table 2. Summary of selected 10 samples used in metagenomic sequencing

Supplementary Table 3. BGI 2.0 adapters sequences

Supplementary Table 4. (A) summary of metagenomic sequencing data generated. The classified reads and unclassified reads were mapped to a curated database using Kraken and Bracken. **(B)** The relative abundance of number of mapped reads to each taxa in curated database per sample.

S**upplementary Table 5.** An overview of OTUs assignment per sample using ITS2 metabarcoding.

Supplementary Table 6. A) Summary of binning details and B) percentage of recruitment for vineyards 4 and 5. Percentage of recruitment summarizes the mean coverage of each split in each bin, and normalize every bin with respect to each other. It is critical to remember that these values do not take the unassembled data into account.

Supplementary Table 1. Details of samples and measured wine parameters.

No.	Sample name	Metabarcoding	Metagenomics	Biological replicate	Vineyard	Sampling date	Alcoholic percentage	Density	Alcohol	Total sugar	Glucose	Fructose	Titrable acid	pH	Volatile acid	Tartaric acid	Lactic acid	Malic acid	Citric acid	Glycerol	Extract	NOPA	NH4
1	w2a_start			2a	2	24/9/2015	0.00	1.095		214.8	104.2	110.6	9.3	3.16	0.055	9.1	-	3.8	0.3	0.6		79	30
2	w2a109	x		2a	2	9/10/2015	0.50	1.088		201.2	101.1	100.1		3.16									
3	w2a112	x		2a	2	12/10/2015	1.11	1.084	8.09	195.4	95.01	95.62	10.07	2.4	-0.01	5.35	0.63	3.55	2.17	-1.64	25.42	98	67
4	w2a114	x		2a	2	14/10/2015	1.13	1.084	8.22	196.1	96.44	95.14	9.93	2.51	0.05	5.46	0.83	3.27	2.33	-1.38	26.12	112	73
5	w2a120	x		2a	2	20/10/2015	1.65	1.079	12.07	180.2	90.29	86.58	13.23	2.3	0.79	6.48	1.05	4.4	2.26	0.16	31.01	117	60
6	w2a_end			2a	2	26/10/2015	3.50	1.066	25.94	149.2	88.57	59.83	13.36	2.28	0.93	6.67	0.87	4.68	1.37	2.62	31.66	133	64
7	w2a_end2			2a	2	15/12/2015	10.73	1.008	83.97	21.3	21.76	2.9	13.81	2.49	1.39	6.94	0.64	4.33	0.31	8.69	32.59	72	6
8	w2bstart			2b	2	24/9/2015	0.00	1.095		214.8	104.2	110.6	9.3	3.16	0.055	9.1	-	3.8	0.3	0.6		79	30
9	w2b109	x		2b	2	9/10/2015	1.30	1.086		201.2	101.1	100.1		3.16									
10	w2b112	x		2b	2	12/10/2015	1.62	1.081	11.85	192.1	96.09	91.2	9.6	2.53	0.08	5.2	0.61	4.06	1.66	-1.43	24.94	93	107
11	w2b114	x		2b	2	14/10/2015	1.99	1.078	14.57	184.9	93.65	86.38	9.74	2.56	0.18	5.62	0.75	4.17	1.16	0.05	25.14	94	110
12	w2b120	x		2b	2	20/10/2015	3.66	1.065	27.1	158.3	87.58	67.27	10.01	2.65	0.41	5.79	0.77	3.91	1.15	2.62	24.28	84	99
13	w2b_end			2b	2	26/10/2015	4.83	1.053	36.22	130.4	82.15	46.94	9.85	2.64	0.46	5.84	0.54	3.98	0.46	4.41	24.44	92	103
14	w2b_end			2b	2	15/12/2015	8.00	1.029	61.36	74.9	75.03	2.97	9.89	2.7	0.57	5.9	0.43	3.75	-0.04	7.03	24.88	104	93
15	w3a_start			3a	3	24/9/2015	0.00	1.091		202.5	97.8	104.7	11.7	3.07	0.082	11.8		3.4	0.3	0.3		68	69
16	w3a109	x	x	3a	3	9/10/2015	1.50	1.074		178.7	88.7	90		2.8									
17	w3a112	x	x	3a	3	12/10/2015	5.65	1.043	42.73	110.7	43.11	63.94	9.73	2.6	0.26	6.18	0.32	3.11	0.37	2.85	21.93	35	12
18	w3a114	x		3a	3	14/10/2015	8.18	1.022	63.21	64	19.72	41.89	9.6	2.66	0.29	6.4	0.22	2.85	0.07	4.61	20.97	16	11
19	w3a120	x		3a	3	20/10/2015	11.41	0.998	90.25	11.4	2.13	8.7	9.47	2.68	0.26	6.13	0.07	2.36	0.18	4.91	20.39	1	4
20	w3a_end			3a	3	26/10/2015	11.60	0.994	92.12	1.1	-0.55	2.25	9.33	2.74	0.25	6.1	0.29	2.59	0.24	5.11	20.17	8	17
21	w3a_end2			3a	3	5/11/2015	11.87	0.993	94.32	0.7	-0.52	2.16	8.86	2.78	0.488	5.03	0.19	2.65	0.35	5.95	20.34	4	7
22	w3b_start			3b	3	24/9/2015	0.00	1.091		202.5	97.8	104.7	11.7	3.07	0.082	11.8		3.4	0.3	0.3		68	69
23	w3b109	x		3b	3	9/10/2015	6.00	1.040		92.41	34.92	57.49		3.07									
24	w3b112	x		3b	3	12/10/2015	9.51	1.013	74.12	44.3	9.92	32.49	9.74	2.66	0.2	6.36	0.21	2.71	0.2	4.79	20.51	4	2
25	w3b114	x		3b	3	14/10/2015	10.64	1.003	83.65	23.4	3.97	18.83	9.62	2.67	0.21	6.26	0.22	2.52	0.17	5.19	20.63	2	2
26	w3b120	x		3b	3	20/10/2015	11.85	0.994	94.01	2	-0.64	3.14	9.51	2.73	0.17	6.05	0.31	2.59	0.24	4.98	20.42	-1	3
27	w3b_end			3b	3	26/10/2015	11.57	0.994	91.91	-0.4	-0.87	1.19	9.2	2.75	0.16	5.92	0.32	2.5	0.23	4.96	20.08	4	13
28	w3b_end2			3b	3	5/11/2015	12.06	0.993	95.8	-0.2	-0.61	1.31	8.97	2.76	0.402	5.02	0.2	2.63	0.32	6.05	20.66	-3	5
29	w4a_start			4a	4	26/9/2015	0.00	1.098		220	106.4	113.6	11.3	3.19	0.386	11.9		3.7	4.4	-0.8		124	0
30	w4a109	x	x	4a	4	9/10/2015	1.00	1.088		201.7	100	101.7	12.09	3.19									
31	w4a112	x	x	4a	4	12/10/2015	4.93	1.054	36.94	133.7	58.06	71.33	12.19	2.29	0.36	7.16	0.38	4.22	0.57	2.47	23.91	15	27
32	w4a114	x	x	4a	4	14/10/2015	8.44	1.026	64.9	72.4	24.11	45.03	12.2	2.38	0.38	7.48	0.21	3.81	0.03	5.1	23.49	2	15
33	w4a120	x	x	4a	4	20/10/2015	12.34	0.995	97.9	2.8	-1.36	3.61	12.03	2.46	0.34	7.36	0.25	3.63	0.14	6.08	22.68	-20	16
34	w4a_end			4a	4	26/10/2015	12.15	0.994	96.46	-0.2	-1.28	1.11	11.68	2.49	0.32	7.14	0.25	3.46	0.19	5.96	22.72	-7	22
35	w4a_end2			4a	4	5/11/2015	12.58	0.993	99.92	0.2	-1.21	1.1	11.2	2.5	0.556	6.06	0.03	3.46	0.26	6.73	22.87	-8	12
36	w4b_start			4b	4	26/9/2015	0.00	1.098		220	106.4	113.6	11.3	3.19	0.386	11.9		3.7	4.4	-0.8		124	0
37	w4b109	x		4b	4	9/10/2015	1.00	1.086		198	99	99		3.19									
38	w4b112	x		4b	4	12/10/2015	1.79	1.080	13.07	188.9	93.34	90.53	11.68	2.41	0.34	6.72	0.85	4.63	1.43	-0.99	25.31	77	98
39	w4b114	x		4b	4	14/10/2015	2.16	1.076	15.83	180.7	90.37	85.94	11.63	2.46	0.45	6.78	0.83	4.38	1.33	-0.11	25.46	70	98
40	w4b120	x		4b	4	20/10/2015	3.47	1.066	25.65	160.4	81.6	73.98	11.86	2.5	0.6	6.83	0.85	4.71	0.96	2.09	25.6	58	94
41	w4b_end			4b	4	26/10/2015	4.00	1.059	29.81	142.3	75.97	63.28	11.46	2.51	0.59	6.86	0.64	4.49	0.67	2.77	25.15	64	96
42	w4b_end2			4b	4	15/12/2015	12.22	0.994	97.03	-0.6	-0.77	0.77	11.69	2.52	0.6	6.81	0.27	3.61	0.18	7.36	23.53	-6	21
43	w5a_start			5a	5	1/10/2015	0.00	1.089		205.3	97.75	102.07	15.8	2.82	0.343	13.8		5.1	0.9	-0.3		55	24
44	w5a109	x		5a	5	9/10/2015	0.22	1.089	1.57	195.5	96.6	98.9	11.33	2.21	-0.12	7.28	0.74	4.05	1.97	-5.09	24.39	102	67
45	w5a112	x		5a	5	12/10/2015	0.87	1.080	6.36	188	89.17	93.33	11.36	2.17	-0.01	6.29	0.45	4.24	1.69	-3	24.07	84	59
46	w5a114	x	x	5a	5	14/10/2015	3.85	1.055	28.78	135.3	54.23	75.95	11.72	2.18	0.13	6.77	0.23	3.8	0.79	1.02	22.98	29	14
47	w5a120	x	x	5a	5	20/10/2015	10.50	1.003	82.66	18.9	1.13	16.39	11.68	2.37	0.25	7.03	0.1	3.3	0.19	5.23	21.74	-3	8
48	w5a_end			5a	5	26/10/2015	11.20	0.995	88.86	0	-1.35	1.24	11.21	2.4	0.23	6.76	0.13	3.15	0.28	4.97	20.97	-3	13
49	w5a_end2			5a	5	5/11/2015	11.58	0.994	91.94	-0.3	-1.15	1.06	10.81	2.46	0.481	5.75	0.02	3.28	0.28	6.03	22.03	0	6
50	w5b_start			5b	5	1/10/2015	0.00	1.089		195.5	96.6	98.9	15.8	2.82	0.343	13.8		5.1	0.9	-0.3		55	24
51	w5b109	x		5b	5	9/10/2015	0.14	1.089	0.98	195.5	96.6	98.9	11.3	2.25	-0.11	7.1	0.81	3.87	2.53	-5.48	24.7	104	63
52	w5b112	x		5b	5	12/10/2015	1.49	1.077	10.94	182.3	87.39	89.47	11.35	2.21	0.05	6.46	0.43	4.28	1.52	-2.41	23.08	76	57
53	w5b112	x	x	5b	5	14/10/2015	4.86	1.049	36.53	122	50.73	67.04	11.51	2.24	0.19	6.9	0.23	3.72	0.49	1.66	22.56	20	19
54	w5b120	x	x	5b	5	20/10/2015	10.34	1.006	81.13	28.5	3.17	23.51	11.53	2.4	0.28	7.15	0.12	3.34	0.09	4.73	21.12	-7	10
55	w5b_end			5b	5	26/10/2015	11.20	0.994	88.83	0.8	-1.33	2.04	10.83	2.47	0.25	6.62	0.11	3.1	0.23	4.47	20.07	-6	18
56	w5b_end2			5b	5	5/11/2015	12.08	0.994	95.88	0.9	-0.98	1.95	10.69	2.46	0.496	5.58	-0.04	3.21	0.39	5.92	21.47	-4	3

Supplementary Table 2. Summary of selected 10 samples used in metagenomic sequencing

Sample name	Sample ID	Vineyard	Biological replicate	Sampling date	DNA input (ng)	Index PCR cycles
w3a109	W1	3	a	9/10/2015	2,01	25
w3a112	W2	3	a	12/10/2015	4,38	20
w4a109	W3	4	a	9/10/2015	< 0.02	20
w4a112	W4	4	a	12/10/2015	5,49	25
w4a114	W5	4	a	14/10/2015	6,60	25
w4a120	W6	4	a	20/10/2015	< 0.02	25
w5a112	W7	5	a	12/10/2015	< 0.02	20
w5a114	W8	5	a	14/10/2015	1,59	20
w5b112	W9	5	b	12/10/2015	3,54	27
w5b114	W10	5	b	14/10/2015	5,58	30

Supplementary Table 3. BGI 2.0 adapters sequences

BGI 2.0 adapter	Sequence (5' -> 3')
AD2_Long_2.0	GAACGACATGGCTACGATCCGACTT
AD2_Short_2.0	AAGTCGGATCGT-[C3spacer]
AD1_Long_2.0	TTGTCTTCCTAAGACCGCTTGGCCTCCGACTT
AD1_Short_2.0	AAGTCGGAGGCC-[C3spacer]

Supplementary Table 4a. A summary of metagenomic sequencing data generated. The classified reads and unclassified reads were mapped to a curated database using Kraken and Bracken.

Sample ID	Number of raw reads	Number of reads after adapters removal	Classified reads (%)	Unclassified reads (%)
W1	142920623	140546873	41,47	58,53
W2	102176062	100461732	73,62	26,38
W3	142621612	139924300	15,33	84,67
W4	189804733	186302461	77,69	22,31
W5	278295910	270670195	65,69	34,31
W6	173458829	170202669	34,33	65,67
W7	142743168	139728638	11,14	88,86
W8	101573197	99306540	63,72	36,28
W9	152459270	150066214	34,98	65,02
W10	188277765	184414440	50,54	49,46

Supplementary Table 4b. The relative abundance of number of mapped reads to each taxa in curated database per sample.

Aspergillus clavatus	0,00021	0,00006	0,00057	0,00004	0,00011	0,0003	0,00113	0,00012	0,00029	0,00021
Aspergillus flavus	0,00027	0,00013	0,00252	0,00005	0,00039	0,00127	0,00405	0,00011	0,00073	0,0007
Aspergillus fumigatus	0,00011	0,00001	0,00011	0,00001	0,00002	0,00005	0,00023	0,00003	0,00007	0,00005
Aspergillus nidulans	0,00006	0,00001	0,00014	0,00002	0,00003	0,00006	0,00024	0,00004	0,0001	0,00007
Aspergillus niger	0,00026	0,00005	0,00072	0,00005	0,0001	0,00028	0,0013	0,00017	0,00049	0,00036
Aspergillus terreus	0,00005	0,00002	0,00016	0,00002	0,00003	0,00005	0,00034	0,00009	0,00012	0,00006
Aureobasidium pullulans	0,00025	0,00012	0,00108	0,00069	0,00068	0,00052	0,00164	0,00096	0,00152	0,00104
Bacillus cereus	0,00002	0,00002	0,00008	0,00001	0,00002	0,00006	0,00023	0,00001	0,00007	0,00009
Bacillus subtilis	0,00009	0,00001	0,00015	0,00001	0,00003	0,00006	0,00018	0,00001	0,0001	0,00003
Bacillus thuringiensis	0,00004	0,00001	0,0001	0,00001	0,00003	0,00005	0,00008	0,00001	0,00005	0,00004
Baudoinia panamericana	0,00008	0,00003	0,00019	0,00006	0,00007	0,0001	0,00049	0,00016	0,00023	0,00013
Bifidobacterium bifidum	0,00003	0,00001	0,00023	0,00001	0,00004	0,0002	0,00049	0,00001	0,00005	0,00004
Botrytis cinerea	0,00159	0,00058	0,00619	0,00479	0,00411	0,0013	0,0068	0,00232	0,02168	0,00377
Brettanomyces bruxellensis	0,0004	0,00008	0,00166	0,00008	0,00006	0,00011	0,00167	0,00019	0,00075	0,00028
Burkholderia cenocepacia	0,00118	0,00032	0,0044	0,00036	0,00034	0,00238	0,00726	0,00057	0,00132	0,00077
Burkholderia pseudomallei	0,00068	0,0002	0,00268	0,0002	0,00018	0,00142	0,00431	0,00033	0,00076	0,00047
Burkholderia thailandensis	0,00065	0,00019	0,00257	0,00019	0,00018	0,00133	0,0041	0,00033	0,00074	0,00046
Candida albicans	0,00067	0,0001	0,00224	0,00007	0,00008	0,00015	0,00214	0,00023	0,001	0,00035
Candida glabrata	0,00103	0,0004	0,00456	0,00037	0,0003	0,00037	0,00094	0,00022	0,00129	0,00059
Citrobacter freundii	0,00054	0,0001	0,00077	0,00007	0,00008	0,00043	0,00155	0,00011	0,00029	0,00016
Clavispora lusitaniae	0,00071	0,00015	0,00129	0,00007	0,00006	0,00011	0,03303	0,00296	0,00482	0,00098

Cryptococcus gattii VGI	0,00066	0,00011	0,00025	0,00002	0,00003	0,00008	0,00032	0,00004	0,00012	0,00006
Cryptococcus neoformans	0,00013	0,00002	0,00009	0,00002	0,00003	0,00005	0,00016	0,00001	0,00005	0,00002
Curtobacterium ammonii genes	0,00003	0,00001	0,00018	0,00001	0,00002	0,00019	0,00031	0,00002	0,00006	0,00004
Curtobacterium flaccumfaciens	0,00008	0,00003	0,00045	0,00002	0,00005	0,00049	0,00087	0,00004	0,00008	0,00008
Debaryomyces fabryi	0,00199	0,00096	0,02395	0,00039	0,00343	0,01073	0,03611	0,00075	0,0065	0,0053
Debaryomyces hansenii	0,00054	0,00006	0,00311	0,00007	0,00008	0,00011	0,0017	0,00026	0,00117	0,00037
Diaporthe ampelina	0,00014	0,00004	0,00049	0,00005	0,00009	0,00022	0,00082	0,00006	0,00022	0,00014
Dyella japonica	0,00035	0,0001	0,00132	0,0001	0,00009	0,0008	0,00221	0,0002	0,00036	0,00025
Dyella thiooxydans	0,00025	0,00007	0,00097	0,00008	0,00009	0,00057	0,00173	0,00014	0,0003	0,0002
Erwinia billingiae	0,00011	0,00005	0,00018	0,00003	0,00005	0,00011	0,00027	0,00003	0,00005	0,00004
Erysiphe necator	0,00003	0,00001	0,00008	0,00007	0,00008	0,00003	0,00026	0,00026	0,0002	0,00013
Fusarium fujikuroi	0,00006	0,00001	0,00019	0,00001	0,00004	0,0001	0,00026	0,00002	0,00007	0,00005
Fusarium oxysporum	0,00056	0,00022	0,00551	0,00012	0,00091	0,00251	0,0071	0,0002	0,00173	0,00109
Fusarium verticillioides	0,00009	0,00002	0,00027	0,00002	0,00005	0,00009	0,00033	0,00006	0,00013	0,00008
Gilliamella bombi	0,00002	0	0,00001	0	0,00001	0,00001	0,00003	0	0,00001	0
Gilliamella intestini	0,00001	0	0,00002	0	0	0,00001	0,00001	0	0,00001	0
Gilliamella mensalis	0,00001	0	0,00001	0	0	0,00001	0,00003	0	0	0
Gluconobacter oxydans	0,00219	0,00087	0,00011	0,00001	0,00001	0,00015	0,00024	0,00002	0,00004	0,00002
Hanseniaspora guilliermondii	0,00034	0,00004	0,00224	0,00005	0,00002	0,00008	0,00033	0,00005	0,00071	0,00031
Hanseniaspora uvarum	0,08172	0,00999	0,54252	0,01478	0,00698	0,02186	0,07382	0,01144	0,16887	0,06743

Kazachstania servazzii	0,00294	0,00064	0,0225	0,00047	0,00037	0,00056	0,00326	0,00069	0,00428	0,00153
Klebsiella pneumoniae	0,0012	0,0001	0,0005	0,00005	0,00008	0,00041	0,00105	0,00013	0,00023	0,00012
Kluyveromyces aestuarii	0,00035	0,00005	0,00163	0,00005	0,00005	0,00008	0,00047	0,00008	0,00062	0,00029
Kluyveromyces marxianus	0,00073	0,00011	0,00384	0,00012	0,00008	0,00023	0,00099	0,00013	0,00125	0,00051
Kluyveromyces wickerhamii	0,00035	0,00013	0,00148	0,0001	0,0001	0,00008	0,00043	0,00011	0,00048	0,00021
Komagataeibacter xylinus	0,00018	0,00007	0,00049	0,00004	0,00003	0,00025	0,00082	0,00007	0,00012	0,00008
Lachancea kluyveri	0,00082	0,00011	0,00476	0,0001	0,00009	0,0003	0,00128	0,00013	0,00131	0,00053
Lachancea thermotolerans	0,00009	0,00001	0,00034	0,00001	0,00001	0,00002	0,0006	0,00006	0,00018	0,00004
Lachancea waltii	0,00222	0,00115	0,01967	0,00056	0,00205	0,00725	0,02593	0,00202	0,00502	0,00422
Lactobacillus brevis	0,00008	0	0,00001	0	0	0,00001	0,00002	0	0,00002	0,00001
Lactobacillus buchneri	0,00008	0	0,00001	0	0	0,00002	0,00004	0	0,00001	0
Lactobacillus delbrueckii	0,01653	0,00001	0,00009	0,00001	0,00001	0,00008	0,00012	0,00001	0,00003	0,00002
Lactobacillus plantarum	0,00019	0,00001	0,00005	0,00001	0,00001	0,00003	0,00008	0	0,00002	0,00001
Lelliottia amnigena	0,00016	0,00005	0,00014	0,00002	0,00004	0,00011	0,00027	0,00002	0,00006	0,00005
Leuconostoc citreum	0,00016	0,00001	0,00002	0,00001	0,00002	0,00004	0,00009	0,00001	0,00004	0,00003
Leuconostoc mesenteroides	0,00004	0,00001	0,00006	0	0,00002	0,00005	0,00003	0,00001	0,00003	0,00006
Leuconostoc pseudomesenteroides	0,00011	0	0,00002	0	0	0,00002	0,00003	0	0,00001	0,00001
Mesorhizobium amorphae	0,00086	0,00025	0,00294	0,00029	0,00031	0,00151	0,00562	0,00052	0,00106	0,00063

Mesorhizobium ciceri	0,00053	0,00015	0,00171	0,00017	0,00017	0,00089	0,00349	0,00034	0,00062	0,00039
Mesorhizobium japonicum	0,00068	0,00019	0,0024	0,00025	0,00022	0,00118	0,00469	0,00047	0,00075	0,00048
Methylobacterium extorquens	0,00044	0,00012	0,00145	0,00019	0,00016	0,00085	0,00408	0,00048	0,00055	0,00028
Metschnikowia sp. 04-218.3	0,00045	0,00011	0,00111	0,00006	0,00012	0,00023	0,00567	0,0007	0,00129	0,00041
Metschnikowia sp. 13-106.1	0,00169	0,00039	0,00693	0,00034	0,00024	0,00042	0,01758	0,00191	0,00466	0,0014
Meyerozyma guilliermondii	0,00053	0,0001	0,00237	0,00009	0,00006	0,00023	0,00273	0,00039	0,00125	0,00042
Micrococcus luteus	0,00043	0,00018	0,00473	0,00008	0,00072	0,00779	0,00631	0,00017	0,00065	0,00037
Micrococcus lylae	0,00014	0,00004	0,00059	0,00002	0,00007	0,00159	0,00148	0,00004	0,00012	0,00008
Micrococcus terreus	0,00019	0,00013	0,00141	0,00005	0,00018	0,00144	0,00129	0,00011	0,00035	0,00045
Mitsuaria chitosanitabida	0,00194	0,00057	0,00719	0,0005	0,00053	0,00313	0,01093	0,00081	0,0022	0,00138
Mucor circinelloides	0,00055	0,00011	0,0006	0,00003	0,00006	0,00017	0,00085	0,00012	0,00047	0,00014
Oenococcus oeni	0,00003	0	0,00003	0	0	0,00003	0,00006	0	0	0
Pantoea vagans	0,00007	0,00003	0,0001	0,00002	0,00001	0,00006	0,00014	0,00002	0,00004	0,00004
Pediococcus damnosus	0,00067	0,00001	0,00005	0	0,00001	0,00002	0,00005	0	0,00001	0,00001
Pichia kluyveri	0,00086	0,00007	0,00022	0	0,00002	0,00001	0	0	0	0
Pichia kudriavzevii	0,00131	0,00016	0,00406	0,0001	0,00013	0,00039	0,0031	0,00025	0,00126	0,00049
Plasmopara viticola	0,00002	0,00001	0,00009	0,00001	0,00001	0,00004	0,00011	0,00002	0,00003	0,00003
Propionibacterium acidifaciens	0,0001	0,00003	0,00155	0,00002	0,00029	0,00083	0,00311	0,00004	0,00015	0,00014
Propionibacterium cyclohexanicum	0,00006	0,00002	0,00089	0,00001	0,00016	0,00048	0,00165	0,00002	0,0001	0,00007

Propionibacterium freudenreichii	0,00009	0,00003	0,00135	0,00002	0,00025	0,00071	0,00265	0,00002	0,00013	0,00011
Pseudomonas syringae	0,00043	0,00015	0,00069	0,00007	0,00009	0,00038	0,00111	0,00007	0,00015	0,00014
Ralstonia solanacearum	0,004	0,00115	0,01498	0,00129	0,00126	0,00946	0,02536	0,00213	0,00473	0,00298
Ramularia collo-cygni	0,00021	0,00003	0,00012	0,00001	0,00002	0,00007	0,00025	0,00004	0,00006	0,00005
Ramularia endophylla	0,00013	0,00002	0,00029	0,00002	0,00003	0,00004	0,00017	0,00003	0,00016	0,00006
Rhizophagus irregularis	0,00116	0,00038	0,00428	0,0002	0,0007	0,00169	0,00778	0,00045	0,00197	0,00118
Rhodotorula sp. FNED7-22	0,00005	0,00001	0,00007	0,00001	0,00002	0,00003	0,00016	0,00003	0,00004	0,00002
Rhodotorula sp. JG-1b	0,00008	0,00002	0,00024	0,00003	0,00003	0,0001	0,00026	0,00002	0,00008	0,00006
Roseateles aquatilis	0,00361	0,00104	0,01332	0,00096	0,00107	0,00571	0,0198	0,00144	0,00401	0,00255
Roseateles depolymerans	0,00101	0,00029	0,00399	0,00028	0,00029	0,00176	0,00632	0,00045	0,00118	0,00068
Roseateles terrae	0,00142	0,00036	0,00315	0,00033	0,00039	0,00134	0,00526	0,00064	0,00189	0,001
Saccharomyces bayanus	0,00064	0,00049	0,00156	0,00043	0,0004	0,00047	0,00064	0,00027	0,00049	0,00038
Saccharomyces cerevisiae	0,64529	0,7761	0,08492	0,69036	0,689	0,67528	0,19766	0,68259	0,20944	0,53149
Saccharomyces paradoxus	0,00672	0,01111	0,00179	0,01074	0,01261	0,00834	0,0021	0,00921	0,00288	0,00923
Saccharomyces pastorianus	0,14544	0,16182	0,01558	0,14759	0,12952	0,15365	0,03674	0,13104	0,03884	0,09344
Scheffersomyces lignosus	0,00018	0,00002	0,00085	0,00003	0,00002	0,00006	0,00095	0,0001	0,00035	0,00014
Scheffersomyces shehatae	0,00038	0,00006	0,00125	0,00004	0,00003	0,00012	0,00342	0,00046	0,00108	0,00028
Scheffersomyces stipitis	0,00044	0,00005	0,00262	0,00006	0,00004	0,00009	0,00127	0,00016	0,00089	0,00035

Schizosaccharomyces pombe	0,00014	0,00003	0,0007	0,00002	0,00002	0,00007	0,00089	0,00009	0,00033	0,00011
Snodgrassella alvi	0,00003	0,00001	0,00009	0,00001	0,00001	0,00004	0,00013	0,00001	0,00002	0,00002
Sphingomonas melonis	0,00023	0,00008	0,0008	0,00009	0,00011	0,00052	0,00211	0,00024	0,0003	0,00015
Sphingomonas phyllosphaerae	0,00018	0,00006	0,0007	0,00007	0,0001	0,00039	0,00161	0,00018	0,00024	0,00012
Sphingomonas wittichii	0,00014	0,00005	0,00051	0,00005	0,00006	0,00027	0,00132	0,00014	0,00017	0,0001
Staphylococcus aureus	0,00018	0,00003	0,00117	0,00006	0,00012	0,00046	0,00154	0,00007	0,00028	0,00018
Starmerella bacillaris	0,00002	0,00001	0,00002	0	0,00001	0,00002	0,00005	0	0,00001	0,00001
Tatumella morbirosei	0,00002	0,00001	0,00005	0,00001	0,00001	0,00003	0,00008	0,00001	0,00002	0,00001
Tatumella ptyseos	0,00006	0,00004	0,00002	0	0,00001	0,00002	0,00004	0	0,00001	0,00001
Tatumella saanichensis	0,00001	0,00001	0,00002	0,00001	0,00001	0,00001	0,00003	0	0,00001	0
Torulaspora delbrueckii	0,0004	0,00021	0,00111	0,00017	0,00014	0,00022	0,00029	0,00006	0,00026	0,00015
Variovorax boronicumulans	0,00103	0,0003	0,00392	0,00029	0,00032	0,00197	0,00668	0,00052	0,0012	0,0007
Variovorax paradoxus	0,0007	0,00022	0,00267	0,00022	0,00022	0,00147	0,00473	0,00038	0,00083	0,00052
Variovorax soli	0,001	0,0003	0,00375	0,00027	0,0003	0,00188	0,00629	0,00048	0,00117	0,00067
Vitis vinifera	0,037	0,01793	0,10939	0,11309	0,12997	0,04576	0,33003	0,1277	0,47125	0,2444
Weissella halotolerans	0,00001	0	0,00001	0	0	0,00001	0,00001	0	0	0
Weissella koreensis	0,00006	0	0,00001	0,00001	0	0,00001	0,00008	0,00002	0,00005	0,00003
Weissella parameseenteroides	0,00008	0	0,00002	0	0	0,00002	0,00003	0	0,00001	0
Wickerhamomyces anomalus	0,00075	0,00017	0,00472	0,00015	0,00021	0,00055	0,00116	0,00013	0,00133	0,00062

Supplementary Table 5. An overview of OTUs assignment per sample using ITS2 metabarcoding.

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w2a109	w2a112	w2a114	w2a120	w2b109	w2b112	w2b114	w2b120	w3a109	w3a112	w3a114	w3a120	w3b109	w3b112	w3b114	w3b120
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	21158	57301	6268	8846	5	4014	42	32	208245	196340	125230	83971	99058	131669	94482	188340
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	0	0	0	1	0	0	0	1	0	0	0	0	0	15	1
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	0	27	2	0	2	1410	0	0	0	0	0	0	0	0	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	139	63	167	0	0	0	0	0	0	0	0	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	415	1	0	0	0	0	0	0	0	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	0	1	0	0	115	16	283	163	2	0	0	0	0	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
OTU_15	Fungi	Basidiomycota	Microbotryomycet	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	114	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	106	4	0	0	4931	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	1	1	0	0	4261	0	0	0	0	0	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	88	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	1029	2101	399	37215	1895	7401	9129	98644	0	0	0	0	0	0	0	48
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	0	0	0	146	741	490	0	0	0	0	0	0	0	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	707	0	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	454	35	68	5	0	0	0	0	0	0	0	1
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	61	0	0	0	0	0	0	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	1	0	0	1750	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	1251	0	0	46	1	1739	0	5	2	0	0	0	2	3	1
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	341	0	0	1	0	0	0	0	0	0	0	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	1317	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	0	0	7	495	103	531	0	54	0	0	0	0	0	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	178	0	62	347	331	194	571	27	0	0	0	0	0	0	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	1163	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	3	0	0	848	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	171	0	0	21	0	0	6	0	0	0	0	0	0	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	626	0	0	0	2	0	0	0	0	0	1	9
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	845	0	0	0	1	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	10	0	3	4	4	0	4	26	0	0	0	0	8	3	5
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	579	0	0	0	0	0	0	0	0	0	1
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	697	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	1	0	0	475	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	516	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	340	0	1195	1	0	0	0	0	0	0	0	0	0	0	2
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	296	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	375	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	327	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	445	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	236	1	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	250	2	0	38	1154	865	0	5	1	0	0	1	0	0	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	160	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	277	1113	0	1487	214	733	737	321	4	0	1	1	0	1	2	1
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	1	0	0	1124	0	0	0	0	0	0	0	0	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	85	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	2
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

OTU #	Kingdom	Phylum	Class	Order	Family	Genus	OTU assignment	w4a109	w4a112	w4a114	w4a120	w4b109	w4b114	w4b120	w5a109	w5a112	w5a114	w5a120	w5b109	w5b112	w5b114	w5b120
OTU_1	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Saccharomyces	Saccharomyces	47722	108123	70854	52869	16139	112	91	15	27503	245647	132545	1446	116186	152211	235431
OTU_10	Fungi	Ascomycota	Leotiomycetes	Helotiales	Sclerotiniaceae	Botrytis	Botrytis	0	1	0	0	1216	2559	0	2	129	8	0	261	41	6	15
OTU_103	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Trichomeriaceae	NA	Trichomeriaceae	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_11	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Cryptococcaceae	Cryptococcus	Cryptococcus	12	0	0	0	0	0	0	0	0	0	0	0	4	0	0
OTU_114	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Torulaspora	Torulaspora	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_12	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	38	0	1	1	0	0	0	0
OTU_121	Fungi	NA	NA	NA	NA	NA	Fungi OTU_121	0	0	0	0	0	0	0	0	3	0	0	302	0	0	0
OTU_123	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Lachancea	Lachancea	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
OTU_13	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	Hanseniaspora	Hanseniaspora	2	0	0	0	38	332	634	0	2	0	0	1	0	0	0
OTU_133	Fungi	NA	NA	NA	NA	NA	Fungi OTU_133	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_137	Fungi	Basidiomycota	Tremellomycetes	NA	NA	NA	Tremellomycetes	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0
OTU_14	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
OTU_147	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_15	Fungi	Basidiomycota	Microbotryomycet	Microbotryomycetes_ord_Incertae_sedis	Microbotryomycetes_fam_Incertae_sedis	Curvibasidium	Curvibasidium	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0
OTU_16	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
OTU_17	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	48	0	0	0	0	0	0	0	0	2	0	0	0	0	0
OTU_18	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	0	0	0	0	0	0	5412	0	0	0	0	0	0	0	0
OTU_19	Fungi	NA	NA	NA	NA	NA	Fungi OTU_19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_2	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Starmerella	Starmerella	0	0	0	0	0	388	1892	0	0	0	0	0	0	0	0
OTU_20	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Tremellaceae	Cryptococcus	Cryptococcus	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_21	Fungi	NA	NA	NA	NA	NA	Fungi OTU_21	24	0	0	0	287	0	0	262	0	0	0	1627	0	0	0
OTU_22	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	Kazachstania	Kazachstania	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0
OTU_23	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Mycoacia	Mycoacia	0	0	0	0	0	2676	0	0	0	0	0	0	0	0	0
OTU_25	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	1925	0	0	0	0	0	0	0	0	0	0
OTU_26	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
OTU_27	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_28	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleribasidiaceae	Vishniacozyma	Vishniacozyma	0	0	0	0	0	0	0	0	0	0	0	1753	0	0	0
OTU_29	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Stemphylium	Stemphylium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_3	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	20	3	0	5	4827	182	10	3491	13	16	2	19790	97	13	55
OTU_30	Fungi	Basidiomycota	Tremellomycetes	Holtermanniales	Holtermanniales_fam_Incertae_sedis	Holtermanniella	Holtermanniella	0	0	0	0	0	0	0	0	0	0	0	1704	0	0	0
OTU_31	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodotorula	Rhodotorula	0	0	0	0	1104	0	0	0	0	0	0	0	0	0	0
OTU_32	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Saccharomycetales_fam_Incertae_sedis	Candida	Candida	0	0	0	0	1271	0	0	3	0	0	0	0	1	0	0
OTU_33	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	NA	NA	Chaetothyriales	0	0	0	0	1397	0	0	0	0	0	0	0	0	0	0
OTU_34	Fungi	Basidiomycota	Tremellomycetes	Trichosporonales	Trichosporonaceae	Apiotrichum	Apiotrichum	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_35	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	0	0	0	0	0	0	0	1330	0	0	0
OTU_36	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Didymellaceae	Epicoccum	Epicoccum	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
OTU_37	Fungi	NA	NA	NA	NA	NA	Fungi OTU_37	13	0	0	0	0	6	0	0	0	0	1	2379	0	0	0
OTU_38	Fungi	NA	NA	NA	NA	NA	Fungi OTU_38	6	0	1	0	62	429	348	313	1	0	1	1121	3	0	0
OTU_39	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Fomitopsidaceae	Antrodia	Antrodia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_40	Fungi	Basidiomycota	Agaricomycetes	Cantharellales	Cantharellales_fam_Incertae_sedis	Sistotrema	Sistotrema	0	0	0	0	0	1345	0	0	0	0	0	0	0	0	0
OTU_42	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	914	0	0	0	0	0	0	0	0	0	0
OTU_43	Fungi	Ascomycota	Eurotiomycetes	Chaetothyriales	Cyphellophoraceae	Cyphellophora	Cyphellophora	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_45	Fungi	NA	NA	NA	NA	NA	Fungi OTU_45	0	0	0	0	0	0	784	0	0	0	0	0	0	0	0
OTU_46	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Sporidiobolus	Sporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_47	Fungi	NA	NA	NA	NA	NA	Fungi OTU_47	0	0	0	0	166	0	0	0	0	0	0	480	0	0	0
OTU_48	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	NA	NA	Saccharomycetales	0	0	0	0	662	0	0	1	0	0	0	0	0	0	1
OTU_49	Fungi	Ascomycota	Sordariomycetes	Xylariales	Diatrypaceae	Eutypella	Eutypella	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_5	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	Metschnikowia	Metschnikowia	77	0	0	0	0	0	0	8528	6421	1417	3	9744	866	13	46
OTU_50	Fungi	NA	NA	NA	NA	NA	Fungi OTU_50	0	0	0	0	0	0	579	0	0	0	0	0	1	0	0
OTU_52	Fungi	NA	NA	NA	NA	NA	Fungi OTU_52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_53	Fungi	Basidiomycota	Agaricomycetes	Polyporales	Meruliaceae	Irpex	Irpex	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_54	Fungi	Ascomycota	Eurotiomycetes	Eurotiales	Aspergillaceae	Penicillium	Penicillium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_57	Fungi	Ascomycota	Dothideomycetes	NA	NA	NA	Dothideomycetes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_59	Fungi	Mucoromycota	Mucoromycetes	Mucorales	Mucoraceae	Mucor	Mucor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_6	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Rhynchogastremataceae	Papiliotrema	Papiliotrema	0	0	0	0	0	2618	0	0	0	0	0	0	0	0	0
OTU_61	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_63	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_64	Fungi	Basidiomycota	Wallemiomycetes	Wallemiales	Wallemiaceae	Wallemia	Wallemia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_65	Fungi	Basidiomycota	Agaricomycetes	Agaricales	Omphalotaceae	Lentinula	Lentinula	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_66	Fungi	Basidiomycota	Tremellomycetes	Tremellales	Bulleraceae	Bullera	Bullera	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_7	Fungi	Ascomycota	Dothideomycetes	Pleosporales	Pleosporaceae	Alternaria	Alternaria	0	0	0	0	0	0	0	4303	1	0	0	1088	14	5	0
OTU_71	Fungi	NA	NA	NA	NA	NA	Fungi OTU_71	0	0	0	0	0	213	0	0	0	0	0	0	0	0	0
OTU_73	Fungi	Basidiomycota	Microbotryomycet	Sporidiobolales	Sporidiobolaceae	Rhodosporidiobolus	Rhodosporidiobolus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_79	Fungi	NA	NA	NA	NA	NA	Fungi OTU_79	0	0	0	0	0	0	0	1	0	0	0	0	3	0	0
OTU_8	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	1	0	2	5	429	2513	0	1005	4	11	1	2346	34	7	5
OTU_88	Fungi	Ascomycota	Dothideomycetes	Capnodiales	Cladosporiaceae	Cladosporium	Cladosporium	0	0	0	0	0	0	0	0	0	0	0	70	0	0	0
OTU_9	Fungi	Ascomycota	NA	NA	NA	NA	Ascomycota	0	0	0	0	0	0	0	3509	5	5	2	3549	3	0	0
OTU_90	Fungi	Ascomycota	Sordariomycetes	Hypocreales	Cordycipitaceae	Lecanicillium	Lecanicillium	0	0	0	0	0	14	11	0	0	0	0	0	0	0	0
OTU_91	Fungi	NA	NA	NA	NA	NA	Fungi OTU_91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OTU_93	Fungi	Ascomycota	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Zygoascus	Zygoascus	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
OTU_97	Fungi	Ascomycota	Dothideomycetes	Dothideales	NA	NA	Dothideales	0	0	0	0	38	0	0	0	0	0	0	0	0	0	0
OTU_98	Fungi	Basidiomycota	NA	NA	NA	NA	Basidiomycota	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Xylella fastidiosa	0,00004	0,00001	0,00012	0,00001	0,00001	0,00008	0,00021	0,00002	0,00003	0,00003
Yarrowia deformans	0,00012	0,00003	0,00024	0,00002	0,00004	0,00013	0,00059	0,00003	0,00012	0,00007
Yarrowia keelungensis	0,00011	0,00001	0,0001	0,00001	0,00001	0,00003	0,00021	0,00002	0,00005	0,00004
Yarrowia lipolytica	0,0002	0,00009	0,00034	0,00005	0,00011	0,00017	0,00058	0,00008	0,00019	0,00009
Zygosaccharomyces bailii	0,00101	0,00112	0,00033	0,00085	0,00081	0,001	0,00034	0,00005	0,00014	0,00012
Zygosaccharomyces rouxii	0,00041	0,00014	0,00169	0,0001	0,00009	0,00011	0,00044	0,00014	0,00049	0,00026

Supplementary Table 6. A) Summary of binning details and B) percentage of recruitment for vineyards 4 and 5. Percentage of recruitment summarizes the mean coverage of each split in each bin, and normalize every bin with respect to each other. It is critical to remember that these values do not take the unasssambled data into account.

A)					
Vineyard 4					
Bin Taxonomy	**Total Size (Mb)**	**Number of contigs**	**Completeness**	**Redundancy**	**Number of genes identified**
Unknown_Actinobacteria	1,87	319	82,01 %	0,72 %	1969
Unknown_Bacteria	3,56	959	55,40 %	10,07 %	4315
Hanseniaspora	4,68	948	44,58 %	0,00 %	2941
Pelomonas	6,08	545	94,96 %	7,91 %	6194
Saccharomyces	9,78	1316	75,90 %	7,23 %	6562
Vineyard 5					
Bin Taxonomy	**Total Size (Mb)**	**Number of contigs**	**Completeness**	**Redundancy**	**Number of genes identified**
Unknown_Actinobacteria	2,34	61	97,12 %	0,72 %	2199
Bradyrhizobium	8,08	585	77,70 %	9,35 %	8103
Hanseniaspora	9,36	582	78,31 %	1,20 %	5475
Metschnikowia	12,3	2682	54,22 %	6,02 %	8768
Pelomonas	6,08	521	97,84 %	9,35 %	6146
Saccharomyces	11,58	248	86,75 %	7,23 %	6480

B)				
Percentage of recruitment				
Vineyard 4				
Bin Taxonomy	**W3**	**W4**	**W5**	**W6**
Unknown_Actinobacteria	11,8 %	0,4 %	6,1 %	39,5 %
Unknown_Bacteria	3,6 %	1,0 %	1,0 %	0,0 %
Hanseniaspora	72,0 %	8,2 %	5,6 %	10,9 %
Pelomonas	7,6 %	1,9 %	1,8 %	10,6 %
Saccharomyces	3,1 %	83,3 %	79,8 %	23,8 %
Others*	1,9 %	5,2 %	5,7 %	15,2 %
Total	100,0 %	100,0 %	100,0 %	100,0 %
Vineyard 5				
Bin Taxonomy	**W9**	**W10**		
Unknown_Actinobacteria	1,3 %	1,2 %		
Bradyrhizobium	1,7 %	0,6 %		
Hanseniaspora	44,2 %	22,8 %		
Metschnikowia	18,1 %	3,9 %		
Pelomonas	6,1 %	4,6 %		
Saccharomyces	22,6 %	61,6 %		
Others*	6,1 %	5,2 %		
Total	100,0 %	100,0 %		

*: Bins with lower 40% completeness

Chapter 5

Sensory relevant chemical interactions revealed with Riesling wines fermented with different microbial communities

Sensory relevant chemical interactions revealed with Riesling wines fermented with different microbial communities

Kimmo Sirén[1,2*], Inês Oliveira[3], Vicente Ferreira[3], Ulrich Fischer[1]

1. Institute for Viticulture and Oenology, Dienstleistungszentrum Ländlicher Raum (DLR) Rheinpfalz, Neustadt an der Weinstraße, Germany
2. Department of Chemistry, University of Kaiserslautern, Kaiserslautern, Germany
3. Laboratory for Flavor Analysis and Enology, Department of Analytical Chemistry, University of Zaragoza, Spain

Highlights

- Riesling wine fermented spontaneously with microbial communities originating from vineyard and cellar.
- Chemical interactions and sensory effects were studied over two consecutive years.
- Vineyard microbial communities effect to aroma and sensory were found important.
- Wine chemical and sensorial variability increased with the vineyard-winery microbial interactions.

Abstract

Wine fermentations relate to controlling of microbial community activities. Microbial compositions of wines play a crucial role in quality although less is known about their interactions. Using targeted chemical and sensory analyses this interaction effect of native microbial populations originating both from the vineyard and in the cellar were studied. Riesling wines were fermented spontaneously with microbial communities originating both from vineyard and from both winery and vineyard. The fermentation location was found to play huge role in influencing wine. The combined effect of winery and vineyard populations was significantly distinguished from vineyard wines in both vintages on both levels, sensory and aroma composition increasing variability. The interactions of the chemical compounds were found to validate multiple existing theories about aroma interactions. With multiple key odorants found that can influence both positively and negatively to sensory perceptions. Such compositional studies shed light upon understanding about underlying mechanisms around complex interactions that affect both wine quality and typicality.

Keywords

Wine, compositional study, spontaneous fermentation, machine learning, GC-MS, descriptive analysis, sensory analysis, terroir

1 Introduction

In winemaking and especially during wine fermentation, controlling the activities of microbial community is crucial. Common practices to control fermentation include applying sulphur (SO_2) to reduce the microbial impact and inoculating with cultured yeast or bacteria. Without such measures, the wine quality can be impaired due to sluggish fermentation rate as reviewed by Malherbe et al. (2007). This can cause multiple odour issues ranging from reduction to oxidation as reviewed by Kraus et al. (1984). However, even with proper control, especially white wines are known to be susceptible to sluggish fermentation derived problems (Bisson, 1999; Ribereau Gayon, 1999).

Notwithstanding, spontaneous fermentations are regaining attention in lieu of personalization or individualization of winemaking as a natural and artistic activity, but also due to the increase of research about non-conventional yeasts as reviewed by several authors (Belda et al., 2016, 2017a; Padilla et al., 2016; Tempère et al., 2018; Varela, 2016). For instance, non-conventional yeasts are generally seen to increase complexity (Tempère et al., 2018), whereas commercial products of the most important one in wine, *Saccharomyces cerevisiae* (*S. cer*), are found to be relatively genetically similar (Borneman et al., 2016), which might be due to their domestication origin (Gallone et al., 2016). Interestingly *S. cer* has been also proposed to have a regionality aspect despite these genetic similarities (Knight et al., 2015).

Profiling microbes from vineyards has been in the focus of recent wine microbiology (Bokulich et al., 2016; Knight et al., 2015; Mezzasalma et al., 2018; Morrison-Whittle and Goddard, 2018). Typically wine microbes are studied using cultured yeast strains (Ciani et al., 2016; Rossouw et al., 2015; Sadoudi et al., 2012; Wang et al., 2016) Such approach does not allow investigation of microbial communities and their interactions (Amann et al., 1995). These interactions are nonetheless crucial in wine fermentation. For example, capability of *S. cer* to complete alcoholic fermentation is suggested to related to microbial biodiversity where increase in species richness can negatively influence the capability of *S. cer* to dominate (Boynton and Greig, 2016). Furthermore, evidence that winery specific microbial communities have an important role in wine fermentation have accumulated over the years (Bokulich et al., 2013; Ciani et al., 2004; Quan and Eisen, 2018; Varela and Borneman, 2017; Xufre et al., 2006). Interestingly, investigations into

microbial regionality have often looked in either vineyard or winery derived communities, without any focus on the interaction between the two different populations. This link has investigated by profiling the elemental composition where the authors found that both vineyards and wineries impacted the elemental composition of wines (Hopfer et al., 2015).

Next-generation sequencing (NGS) based approaches that target specific gene regions, such as metabarcoding, allow studying wine microbes without strain isolation. These methods have revealed that typicality, terroir and sense of place, features that are found to be important in high quality wines, translate to wine microbial communities (Belda et al., 2017b). For example, certain microbial communities have been linked to specific regions (Bokulich et al., 2014; Morrison-Whittle and Goddard, 2018). In general, microbiome studies have dramatically increased in recent years as reviewed by Belda et al., (2017b) and Stefanini and Cavaliere (2018), highlighting their relevance in wine research. However, because the approaches are relatively new in wine, inter-study comparisons are often impossible due to the lack of standardisation such as in DNA extraction methods (Mak et al., 2019)

While relative certainty about the regionality of microbial communities has been established, functionality of these communities remains unstudied. Few functional analyses have been done (Sirén et al., 2019b), however, their impact on sensory properties is unknown. Therefore, combining multiple ‘omics methods have been proposed to offer insights into these questions in wine research (Sirén et al., 2019a). Such approach could improve the understanding how microbial activities are translated to sensory perceptions. Although sensory links remain unestablished, early studies have made vague linkages between microbial communities and chemical compounds. Bokulich et al. (2016) found impact of fermentative yeasts and bacteria on relative abundances of putative compounds whereas another study observed links with certain wine fungal community members and thiols (Knight et al., 2018).

Linking changes in chemical composition to sensory perception is not trivial thus making investigations about perceivable the microbial effects to wine even more complicated. Usually the actual chemical concentrations and sensory analysis have been combined using univariate and multivariate approaches that are based on (multi)linearity (Fischer et

al., 1996; Lee and Noble, 2003, 2006; Noble and Ebeler, 2002; Sáenz-Navajas et al., 2018; Schüttler et al., 2015; Siebert et al., 2018). However, the odour perception of a compound depends on the odour detection threshold (OT), which is influenced by the properties of the compound but also of the matrix (Pineau et al., 2009; Sáenz-Navajas et al., 2010). Therefore, approaches that base their hypothesis on (multi)linear correlations of chemical compositions can be misleading in terms of sensory perceptions. For such reasons, odor activity values (OAV) which relate chemical concentrations to sensorial significance, continue to be studied (Wu et al., 2016). Although both of these concepts, OT and OAV, have drawn several critical comments, due to multiple issues regarding the matrix effect and methodology as briefly reviewed in (Fischer, 2007; Lawless and Heymann, 2010; Wu et al., 2016) they continue to be of use (Juan et al., 2013). Furthermore, it should be kept in mind that OAV was originally designed solely to understand compositional or proportional relationships of different odors in a mixture instead of intensity measurement (Guadagni et al., 1966). Tools for compositional studies can be found in the NGS based methods using relative abundance of microbial compositions. Proportionality is known for its' complicated nature and spurious correlation so multiple approaches have been introduced to overcome these challenges (Egozcue et al., 2003; Reimann et al., 2017; Schwager et al., 2017).

Another issue of linking chemical composition to sensory perception relates to the complex ties and interactions of odors in wine (Campo et al., 2005; Ferreira et al., 2016; San-Juan et al., 2011). Wine's chemical composition would need to be fully investigated but methods allow often only targeting few key odors of interest or putatively identified compounds. It has been suggested that any food odour is a relatively simple mixture of few key volatiles and therefore full characterization of chemical compounds is not needed. (Dunkel et al., 2014). Therefore, linking chemical similarities to odour type continues to be of interest (Chastrette and Rallet, 1998; Nozaki and Nakamoto, 2018) while biological factors also suggested to be relevant (Poivet et al., 2018).

Riesling is known for its multifaceted aroma profile (Fischer, 2007; Rapp and Mandery, 1986), as well as capability to showcase typicity and sense of place (Bauer et al., 2011; Schüttler et al., 2015; Winton et al., 1975). By being able to target multiple of its key aroma compounds, we chose to investigate Riesling and determine if there is any sensorial impact originating from the interaction of the most common microbial

communities affecting the winemaking: the vineyard and the cellar microbial communities. In the holistic approach, we focus on the idea that wine related odor is composed of complex interactions between aroma and non-volatile compounds. We chose to investigate the compositional volatile matrix using Bayesian statistics as well as non-linearity dependent machine learning approaches. In order to relate sensory associations of compound abundances to understanding of what is happening during wine fermentation, our investigations include chemical similarities based on "chemical taxonomy" (Djoumbou Feunang et al., 2016), known precursors sources (Dunkel et al., 2014) and established chemical-odour relations (Wishart, 2014). By using a plethora of novel approaches to wine studies, this study shed light upon the chemical compound interactions and their sensory relevance, additionally finding that sensorial and chemical compositions were heavily influenced by the origin of the alcoholic fermentation, vineyard and the year.

2 Material and Methods

2.1 Trial set up and sample collection

Riesling grapes were harvested from five vineyards in the Pfalz region in Germany in 2015, which were extended to seven in 2016 (Table 1). Briefly, the grapes from each vineyard site were split in two different processes, with one part subjected to a standardized winemaking protocol described below, and the other portion undergoing customary winemaking at the respective cooperating wine estates, similar to previous study (Bauer et al., 2011).

For the customary winemaking, the grapes underwent a conventional winemaking scheme in the corresponding wine estates' cellars without any inoculation of cultured yeasts. When fermentation was complete (below 0° Oechsle / density below 1), samples were collected in 34 L carboys and transported to the pilot plant at the Institute for Viticulture and Oenology (DLR Rheinpfalz, Neustadt an der Weinstraße, Germany) for standardized post-processing.

For the standardised winemaking protocol 250 kg of grapes were individually handpicked

using cloves and sterilised shears and samples into autoclaved sterile flat plastic bags (neoLab) holding 10 kg each from each vineyard. The bags were sealed inside the vineyards to avoid any contamination from outside the vineyard. For vineyard 1 in 2016, 700 kg was processed. Pilot-scale vinification took place in a temperature-controlled intermodal container set at 20 °C and all equipment was sterilized by using ozonated water (3%) (Univog 500, Röha),. Whole cluster pressing (Europress 15, Scharfenberger Maschinenbau) was followed by overnight sedimentation, after which the must was transferred by gravity into three full 34 L carboys per each vineyard sealed with silicone bungs and airlocks, thus considered as fermentation triplicates. At onset of the spontaneous fermentation, 5 L of must was transferred by gravity out of the carboy to avoid overflowing. Fermentations were considered as finished with less than 9 g/L of total sugars, which is the legal limit for dry wines in Germany. The wines were then racked by gravity to 25 L carboys. The carboys were subsequently transported to storage at +4 °C until bottling. Carboys were kept topped up all times and using glass marbles to maintain a complete filling. SO_2 was first added two months after each racking of the finished ferments, using a 18% liquid potassium hydrogen sulphite solution (La Littorale P18, Erbslöh). Initial SO_2 dosing was 50 mg/L in 2015 and 60 mg/L in 2016. Wines were bottled with free SO_2 25 -30 mg/L in the end of March of the following year. The bottling was done by first racking the wine by gravity into carbon dioxide (CO_2) flushed beer kegs, and then bottled to CO_2 flushed screw capped 7500 ml bottles.

Furthermore, in order to compare differences arising from conventional way of carrying out fermentations, such as the use of commercial yeast cultures and use of sulphur, between native fermentations and standard winemaking, smaller side studies were carried out in 2016. For the inoculation study, must fermented in the winery 7 (wgt7_t209) was inoculated with commercial yeast 1 (SIHAFERM™ Element, Eaton) and must from vineyard 1 (wb1) was inoculated with two commercial yeast products in fermentation duplicates, yeast 1 (17a,b) (SIHAFERM™ Element, Eaton) and yeast 2 (18a,b) (Fermivin® 4F9, Oenobrands). The following vinification workflow was the same as described above. To study the use of sulphur, two doses of SO_2, 40 mg/L (15a-c) and 80 mg/L (16a-c) were added to part of the must (equivalent to 6 carboys) from vineyard 1 during racking. Bottled chemical parameters of the samples are found in Table 1.

2.2 Chemical analysis

The wines' chemical composition was analysed in the following late spring of the vintage in the Laboratory for Flavor Analysis and Enology, University of Zaragoza, Spain. In order to evaluate the wines with the same "physical" age, targeted chemical analyses were performed using GC-MS as previously described for major (Ortega et al., 2001) and minor compounds (López et al., 2002), as well as for aldehydes (Bueno et al., 2014) and thiols (Mateo-Vivaracho et al., 2010). Additionally, classical wine parameters (ethanol, residual sugar, total acidity, volatile acidity, and pH) were determined using Fourier transform mid-infrared spectroscopy with the white wine calibration provided by the manufacturer (Foss WineScan FT120). Free and total SO2 were analysed with enzymatic test kits (REF984634 and REF984345; Thermo Fisher Scientific; Konelab 20 analyzer, Thermo Fisher Scientific).

2.3 Sensory analysis

The wines' sensory properties were investigated using descriptive sensory analysis with two panels, in the late spring of the following year, in order to evaluate the wines at their respective same "physical" age. The panels consisted of 16 panelists (nine male, seven female, aged between 24 and 60 years) for 2015 wines and 20 panelists (eleven male, nine female, aged between 25 and 61 years) for 2016 wines. In both panels, the panelists consisted of external participants and staff members from the Institute for Viticulture and Oenology. All panelists were considered experienced and had previously participated in descriptive analysis sessions.

Training sessions consisted of similar approach as previously described by Sherman et al. (2018). They were carried out using subsets of the wines from the study to select attributes and build the consensus among the standards (Table S1) over 1-hour sessions. For 2015 wines, four training sessions were done and five for the 2016 wines. The panelists were not informed about the specific context of the study and generated the attributes by consensus similar to (Siebert et al., 2018). The reference sensory standards were built from a reference wine which was also served as a warm-up for each session. The reference wine was a neutral dry Riesling base wine intended for sparkling wine

production. The initial profiling of the wines was done in the first session using napping with forced grouping and free-choice descriptors (Louw et al., 2013).

During sessions, 30 mL of wines were served at 14 °C in plastic-lid covered black DIN 10960 wine-tasting glasses (Schott Zwiesel, EAN 4001836016803). Before each session, panelists were required to match unknown reference sensory standards to a corresponding descriptors for evaluating panel performance and serving as a warm-up. Three-digit coded samples were presented in a randomized order for among assessors and between sessions. For 2015, the tastings consisted of duplicate tastings of the biological replicates totaling 20 wines over 3 sessions, while for 2016, duplicate tastings were carried out for 36 wines over 12 sessions. Panelists rated the attribute intensities of the wines on 10-point unstructured scales anchored by “weak” and “intense” using FIZZ environment (Version 2.50, Biosystemes). Odor intensities were evaluated within a session in a comparative way, taste and mouth-feel attributes in monadic order.

2.4 Statistical analysis

2.4.1 Chemical metadata acquisition

Automated structural classification of chemical entities (Djoumbou Feunang et al., 2016) was carried out using the ClassyFire batch and IUPAC International Chemical Identifiers (Heller et al., 2015). FoodDB database was used to infer the corresponding flavor groups for each compound and with manual addition of 2-methylpropan-1-ol, based on the web-version as “wine-like”. Acquired metadata can be found from supporting information (Table S2).

According to the work of Dunkel et al. (2014), the measured compounds were divided into four groups (generalist, individualist and intermediaries). “Not found” category was added for compounds not classified in the original work. The information about biosynthetic precursors were also utilized, assigning the measured compound precursors to originate from carbohydrates, amino acids, isoprenoids, polyphenols or unsaturated fatty acids (Dunkel et al., 2014). The precursors for compounds not found were used following: phenylmethanol as 2-phenylethanol; methyl vanillinate, ethyl vanillinate and acetovanillone as 4-hydroxy-3-methoxybenzaldehyde; citronellol as linalool and geraniol; 2-methylpropyl ethanoate as 3-methylbutyl acetate; ethyl dihydrocinnamate as ethyl

cinnamate; alpha-ionone as beta-ionone; ethyl propanoate and as ethyl octanoate;benzaldehyde and ethyl lactate and oxolan-2-one and diethyl succinate as empty.

2.4.2 Chemical relationships

Seventy five compounds were targeted and measured for 56 individual wines (Table S3). The missing values for each compound were replaced with a random value between zero and the limit of detection, or mean values when not measured for the whole set.

The wine aroma was treated as compositional data, in two ways: the co-occurring relative abundance of features together and the odour relative compositional correlation. Essentially, the chemical concentrations were analysed as relative values or transformed to OAV-values using values derived existing literature (Table S2) followed by transforming features in both sets of samples to relative proportions. In the co-occurring relative approach, each compound was scaled between 0 and 1 and each sample was divided by the sum of values. In the OAV approach, the compounds were transformed to corresponding OAVs (see the original methods (Bueno et al., 2014; López et al., 2002; Mateo-Vivaracho et al., 2010; Ortega et al., 2001) for corresponding compound threshold references) as one entity where the total value of the aroma contributions are to one (Guadagni et al., 1966), and finally each sample was divided by the sum of value for both approaches.

Preliminary inspection was done using the using affinity propagation (Bodenhofer et al., 2011; Frey and Dueck, 2007). In order to detect pairwise associations Bayesian method, BanOCC (Schwager et al., 2017) was applied for analysing covariance of the compositional relative abundance data. The approach was done in using RStan v2.18.2 (Stan Development Team, 2018) in R v3.4.0 using the banocc package v1.2.0 (Schwager, 2015) adjusting the prior and shrinkage parameters as follows, n = 0, L = 30, a = 0.5, b = 5, max_treedepth = 15, adapt_delta = 0.99. Six Markov chains were equipped with 20000 iterations including 4000 warm-up iterations for each chain. The iteration length was estimated successful based on generating Rhat values under 1.1.

Following the computation, 95% highest posterior density (HPD) interval was computed and further used to select features for cluster analysis. Hierarchical agglomerative clustering was applied on the median of the HPD estimates. The clustering was done

using complete method on the distance transformed correlation matrix using Scipy v1.1.0 (Jones et al., 2001) and Seaborn v0.9.0 (Waskom et al., 2017).

For further posterior inference, the 95% HPD interval was used to find the highly correlated features (nodes) with the pairwise relationship (edges) selected that their interval did not include values between -0.15 and 0.15. These edges were retained for network analysis done using NetworkX v2.2 (Hagberg et al., 2008). Betweenness centrality (Freeman, 1977) was computed in Networkx custom algorithm (Brandes, 2001).

2.4.3 Feature selection and feature importances for chemical classification

Supervised learning approach was used in order to find features that separated the samples in three different levels: The samples were classed among the two years (2015 and 2016), two locations of fermentation: winery ('wgt') and vineyard ('wb'), and finally the individual vineyard (vineyard 1-7) effect across the others. The feature selection algorithm Boruta v6.0.0 (Kursa et al., 2010), was applied separately using random forests (Wright and Ziegler, 2017) for chemical analysis data by generating 3 different results. The parameters were set as 5000 trees and 500 iterations, the confidence level was set to 0.01 adjusted with Bonferroni correction. In addition to visualise the impact of feature selections, scikit-learn v0.19.1 (Pedregosa et al., 2011) was used to decompose the selected features and scaled (between zero and one) before principal components analysis (PCA).

2.4.4 Sensory evaluation

In the generic descriptive analysis, the panel reproducibility and agreement was done following the suggested workflow (Naes et al., 2010). Sensory panel performances were checked using analysis of variance derived from *panelperf* and *paneliperf* functions from SensoMineR (Le and Husson, 2008).

Before more in depth analysis, the raw data from both years were analysed separately and first decomposed using a constrained correspondence analysis (CCA) (Legendre and Legendre, 2012; ter Braak, 1986) using vegan v2.5-3 (Oksanen et al., 2018). The interaction between repetition and wines during sessions was constrained with partialling out the panelist effect. The significance of axes was evaluated with 500 permutations

inside the panelist strata (Legendre et al., 2011).

The individual features of each judge against the above described classification tasks of samples in three levels, were studied by using the same random forest feature selection wrapper with parameters as described above, however the location effect was tried to be constrained when analysing the vineyard. During the vineyard feature selection task, the location group means were subtracted from both groups for variables found important for location Additionally the judge descriptor effects on mean differences were investigated using univariate independence test as performed with aggregation and permutation strategies using coin -package (Hothorn et al., 2006). The resulting mean values were retained using discovery rate 5%.

All the figures were generated using pandas v0.23.4 (McKinney, 2010), Matplotlib v3.0.0 (Hunter, 2007) and Seaborn v0.9.0 (Waskom et al., 2017) libraries followed by further post-processing in Inkscape v0.91 (https://inkscape.org/).

3 Results

All the Riesling wines finished fermentation (Table S1). The wines originating from the winery were generally observed to ferment longer and they additionally experienced malolactic fermentation during the vinification whereas none the vineyard samples did malolactic fermentation (Table S1).

3.1 Sensory analysis validates some of the observations

The panel performances of sensory sessions for both years were investigated using univariate statistics. The main influence came from the judges, followed by the product significances (Table S4).

The raw sensory data was continued to be investigated by decomposing the raw data with a CCA. Permutation test found three of the dimensions significantly distinguishing the individual samples based on the inspection of the products repetition interactions (Table S5). This is also evident when visualizing the scores (Figure 1 and Figure S1). The location effect was the most observable effect as differentiated by the first dimension of the CCA for both years and for 2016 the second dimension also contributed (Figure 1).

This separation was clearly driven by the descriptors, fruity, lemon, "apple.peach", floral, exotic, sour and mouthfeel, for both years (Figure 1). Whereas two other dimensions contributed not only to outliers such as oxidation with one tasting duplicate (wgt 3 in 2016), but also retained the vineyard effect for both years such as vineyard 5 (Figure S1).

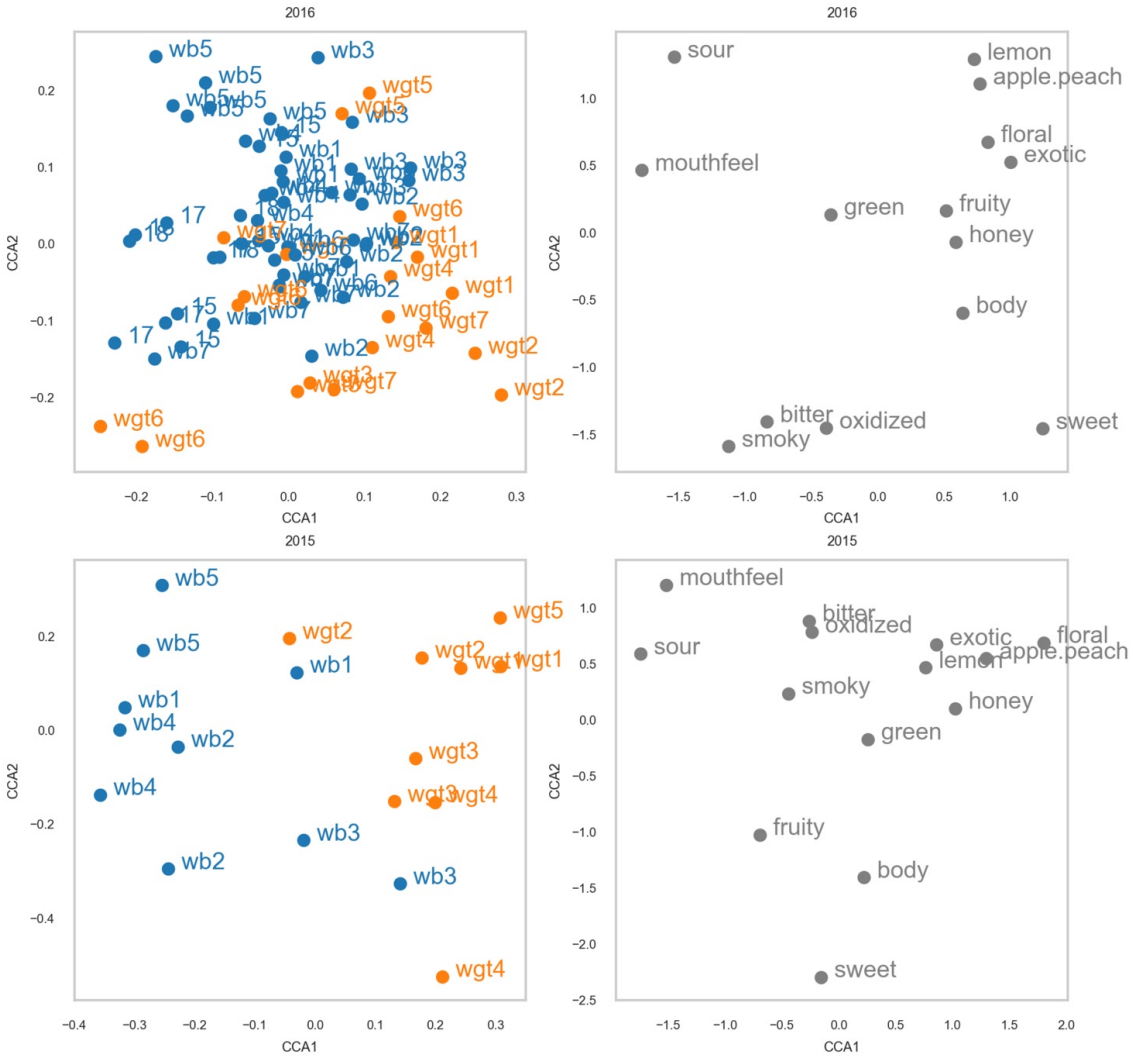


Figure 1. Constrained Canonical Analysis (CCA) decomposition of the sensory analysis raw data with product and tasting replicates as individual samples (left panel) and variable scores (right panel) visualised in the dimensions 1 and 2 for 2015 (lower panel) and 2016 (upper panel) vintages. Labels are composed of 2 sections: wb = vineyard and wb = winery; numbers (1-7) represent each vineyards; 15 = vineyard 1 with SO_2; 17&18 = vineyard 1 with yeast 1 & yeast 2. Blue = pilot-scale fermented, orange = winery fermented.

In order to further investigate the effect of the individual judge, a random forest based feature selection was done. Judges found attributes for location differences in both years were much more frequent than for any of the vineyards (Figure S2). Furthermore, it seemed that the same vineyards lacked distinguishability (vineyard 1 and vineyard 3). Visualizing the chosen attributes separately, it was found that "fruity taste" attribute was differentiating the locations with both years. From a closer inspection of the attributes chosen by the featured selection, the attribute fruity taste was observed to be the driving force of all separations, besides lemon (Figure 2).

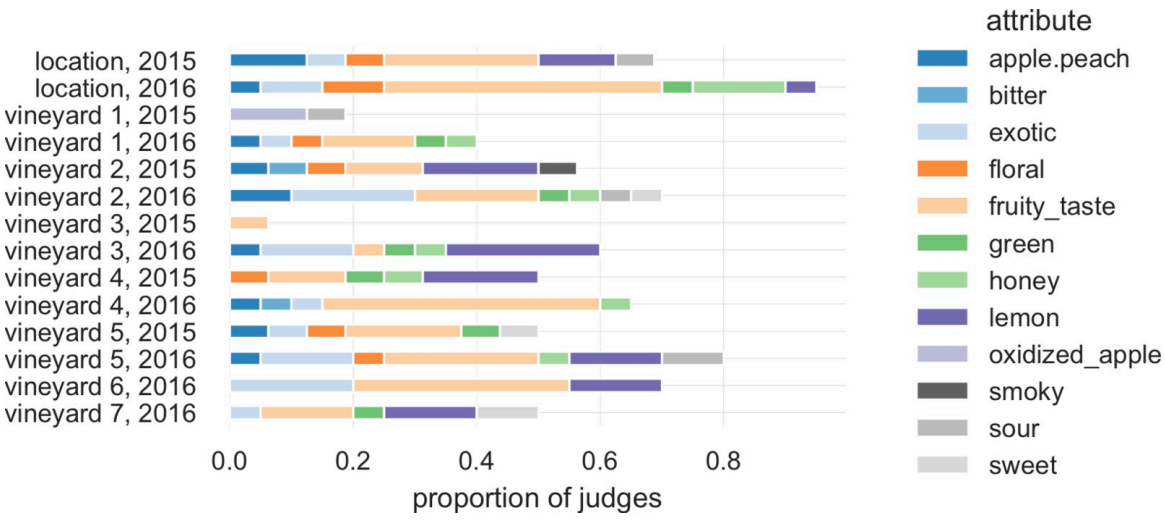


Figure 2. Barplot of stacked random forest selected attributes against the proportion of judges. The judges counts are scaled to corresponding panel sizes (2015: 16 judges, 2016: 20 judges) Location corresponds to the location of fermented wines [wb (vineyard) vs wgt (winery)]. Note that the attribute can mean positive or negative importance.

Furthermore, in order to extract more information from individual judges, investigations to the significant descriptor differences among several classifications were carried out. The mean difference changes that were discovered with permutation established confidence levels under 95%. The winery wines were often considered to have more honey, floral and exotic fruit, whereas the attributes that were found in higher mean difference changes in the vineyard 5 wines than the rest, consisted of floral, apple and exotic (Figure S3).

These three different sensory approaches (CCA, random forest feature selection and mean difference changes) while looking directly at the raw data, derive clearly that various effects such as the fermentation location separation and vineyard 5 separation are the most distinguishing differences. Therefore, the linking the chemical analysis data to find out these differences were further investigated.

3.2 Chemical similarities provide understanding of the matrix effects

Investigating the contribution of chemical compounds we applied two approaches. Firstly, by using feature selection to find the compounds that are differentiating most between the multiple factors: years, fermentation locations and vineyards. Second approach is trying to build a better understanding of the relationships among the chemical compounds.

The first preliminary investigations were done with the affinity propagation based unsupervised clustering which differentiated the samples to three clusters (Figure 3). First cluster consisted of only 2015 wines, the second had the 2016 wines from vineyard 1 and vineyard 5. This included the sulphur experimented and both yeast inoculated wines. The third cluster was composed of the rest of 2016 wines. This clear separation was influenced by 2-phenylacetaldehyde and beta-damascenone, exerting a strong leverage effect. Their concentrations differentiated the years but may also play a key role in sensory perception while differentiating the two years based on the total abundance generated by summing all the OAV values for each wine together (Figure 4). Furthermore, in 2016, another outlier cluster expressing heavy 2-phenylacetaldehyde values were observed from the inoculated yeasts, as well as the high acid low pH wines from vineyards 4 and 5.

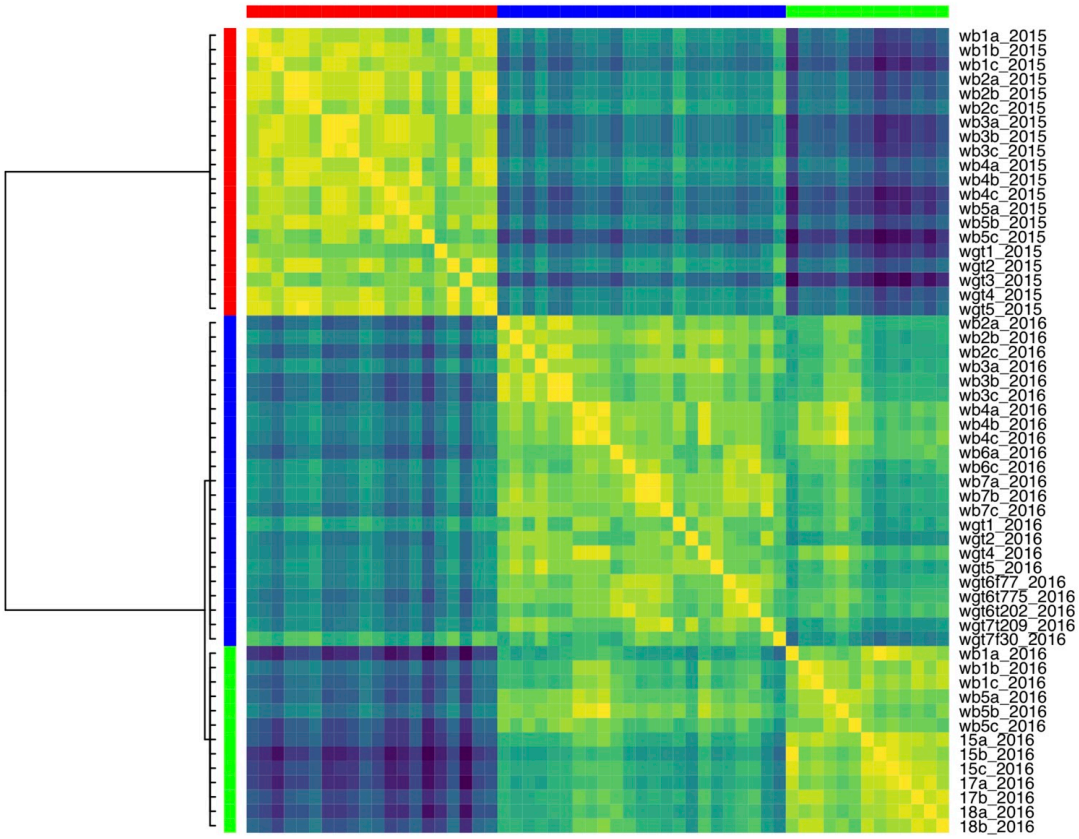


Figure 3. Affinity propagation based sample-sample clustering established three clusters of wines. red: cluster 1 (2015 wines); blue: cluster 2 (2016 wines) 2; green: cluster 3 (2016 wines from vineyard 1 and wb5).

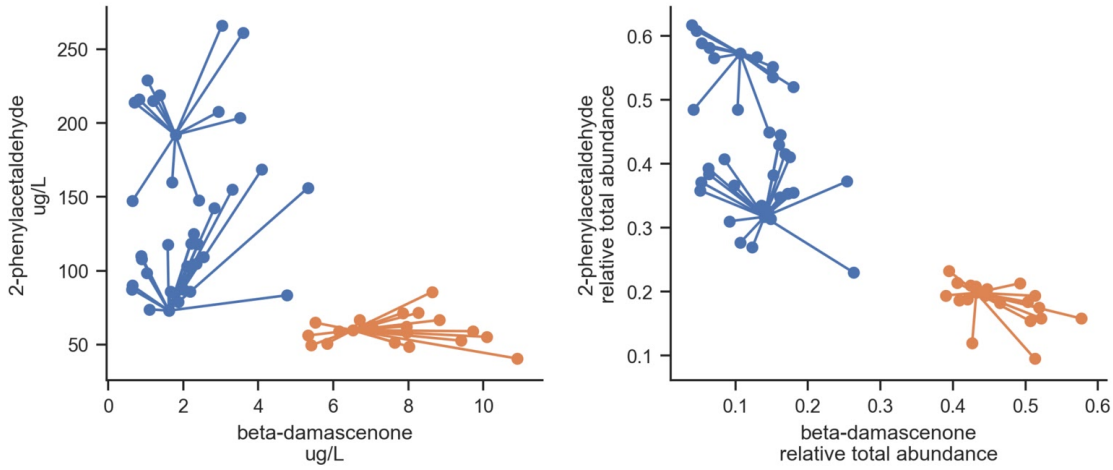
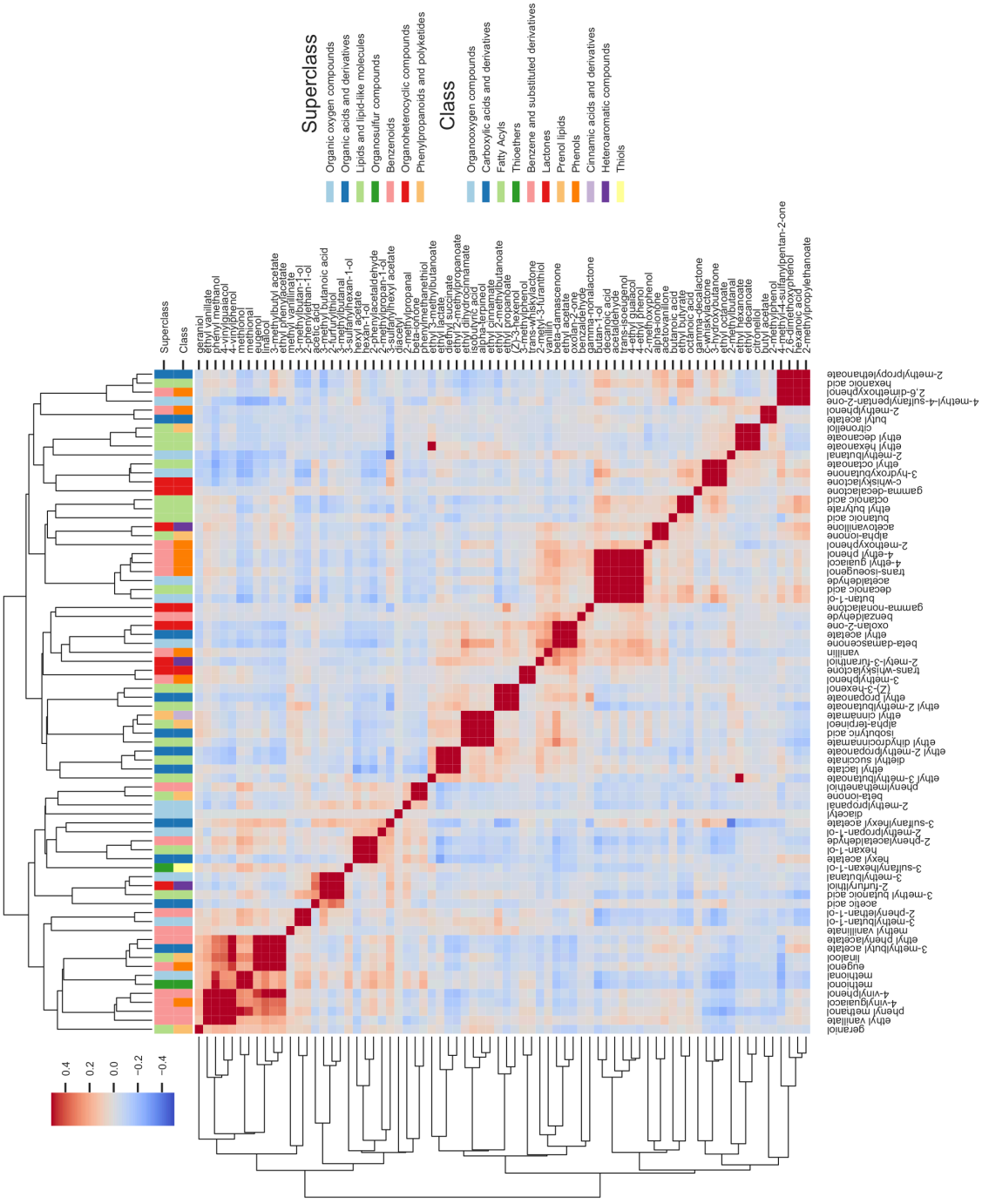


Figure 4. Compositional behaviour shows differences between two years, which might be even more extremely related to sensory differences. The raw values (left) and OAV (right) transformed values and matched to total relative abundance. Blue: 2016 wines and orange: 2015 wines.

In order to find and estimate the importance of the chemical compounds, multiple feature selections were performed. An estimation of the effect of the accepted variables' importances was done following by decomposing the chosen features for visualization with added bivariate gaussian density estimation for each single classes (Figure S4) As the data was found to be significantly affected by the year differences with 50 of all variables selected as important between 2015 and 2016, the follow-up strategy was to process the features as individual years. For 2015 wines, 19 volatile compounds were accepted for their importance to distinguish between the two fermentation locations and 29 compounds for the vineyard separations. For 2016 wines, 23 compounds were found important for fermentation location, whereas 34 for vineyards. Overall, more features were found to be important with individual vineyard often showcasing unique levels of a compound (Figure S5). It is clear that the fermentation locations play a huge role, where most of the compounds tended to be found in higher concentrations in the winery ferments and the variance originating from location could not be constrained but stayed evident decomposed visualization of the vineyard variables (Figures S4 and S5).

Additionally the interactions of the chemical compounds were studied. We investigated the chemical compound correlations in two levels: the co-occurring relative abundance of features together with and the odour relative compositional correlation. The co-occuring of features was investigated by clustering the distance transformed results from the



Bayesian compositional correlation analysis generated correlation matrix of median estimates with 95% probability (Figure 5). Two main groups originate from the agglomerative clustering, both of which can be divided further into several subgroups (Table 2). The division seems to be related to both chemical taxonomy classification but also to precursor content. Group 1 consists of 28 compounds of which 15 can originate

from amino acid precursors (Figure S6), but also the most abundant generalist aroma compounds found previously for all food items in general (Dunkel et al., 2014). This group can be further divided to two subgroups. The first subgroup consists of benzene derivatives, alcohols and their acetates but also sulphur compounds: methionol, methional, linalool, geraniol, 3-methylbutyl acetate, ethyl phenylacetate, 3-methylbutan-1-ol, 2-phenylethan-1-ol, phenyl methanol, eugenol, ethyl vanillate, methyl vanillinate, 4-vinylguaiacol and 4-vinylphenol. The second subgroup consists of straight-chain carboxylic acids and the corresponding aldehydes, together with thiols and phenolic acetals: hexan-1-ol, hexyl acetate, 3-sulfanylhexan-1-ol, 3-sulfanyl hexyl acetate, 2-methylpropanal, 2-methylpropan-1-ol, 3-methylbutanal and 2-phenylacetaldehyde.

Group 2 consists of two subgroups which can both be divided to additional clusters (Table 2). The first subgroup consists of two subclusters. The first subcluster includes mainly ethyl esters of organic acid and lipid-like molecules as well as cinnamic acid derivatives. Together with the previously described groups, most of the amino acids are in these 3 groups. The second subcluster is a complex group consisting of lactones and aromatic compounds but also ethyl acetate. Although their impact on aroma is not clear, these group of compounds are found in influencing OAV levels and constitute to the balsamic odour group as well the fruity characters (Figure S6).

The second subgroup is composed of 3 subclusters. The first on of the subclusters is the most well-defined group and consists of compounds have been made by reactions during the vinification: butan-1-ol, decanoic acid, acetaldehyde and three phenols: trans-isoeugenol, 4-ethyl guaiacol and 4-ethyl phenol. Last two subclusters include all the compounds that can originate from fatty acid precursors besides the previously found c_6-alcohols and acetic acid. These subclusters are mainly composed of fatty acids and their acetates. For instance, one cluster only contains carboxylic acids and ketones and a related compound, whereas the final group contains compounds which are esters and/or have carbonyl structures.

Figure 5. Clustered heatmap of chemical similarities based on relative abundance of each wine and Bayesian compositional correlation analysis. Clustering using 'complete' method was done on the median estimates transformed to distances. Superclass and class correspond to classyfire (Djoumbou Feunang et al., 2016) taxonomy annotations.

Followingly, we investigated the the links between OAV values. More negative relationships are shown. Two main groups originated from the agglomerative clustering as well and in general more structure is evident (Figure 6). However, instead of looking at group behaviour, we further went on to investigate pairwise relationships.

3.3 Network analysis uncovers links

We further investigated the pairwise relationships between the classifications. The pairwise relationships, edges, that had either a higher than 0.15 rho or lower than -015 rho within 95% probability resulted to 117 edges for OAV set and 77 edges for the scaled raw compounds (Tables S6 and S7). For the raw compounds interactions, the strong positive correlations were observed between multiple pairs (Table S6). Examples of median values of positive correlation pairs include 3-methylbutyl acetate and ethyl phenylacetate (rho = 0.86); alpha-terpineol and ethyl cinnamate (rho = 0.86); 2-methylpropyl ethanoate and hexanoic acid (rho = 0.86); methional and methionol (rho = 0.84) and citronellol and ethyl decanoate (rho = 0.85). In general, the volatile phenols correlated strongly between trans-isoeugenol, 4-ethyl phenol (4-EP) and 4-ethyl guaiacol (4-EG) (all rhos higher than 0.8). Their potential precursors 4-vinylphenol (4-VP) and 4-vinylguaiacol (4-VG) were strongly correlated together (rho=0.8) albeit no certainty could established between their ethyl derivatives (rho $\in$ [-0.15,0.15]). The only significant negative correlation was established between 2-methylbutanal and 3-sulfanylhexyl acetate (3-SHA) (rho = -0.38).

For the OAV abundance based compounds interactions, the negative correlations were observed between multiple pairs besides the above 2-methylbutanal and 3-SHA (rho = -0.49) (Table S7), albeit significantly below ODT 3-hydroxybutanone (max 0.06) was found to negatively correlate with other compounds: methional (rho = -0.52), methionol (rho = -0.52), 2-phenylethan-1-ol (rho = -0.50) and 2-phenylacetaldehyde (rho = -0.49). Two of the most sensory relevant negative correlations were observed between 3-sulfanylhexan-1-ol and beta-damascenone (rho = -0.53), but also ethyl lactate and hexyl acetate (rho = -0.49). For the positive pairwise correlations, with the raw values, the highest links were between pairs 3-methylbutyl acetate and ethyl phenylacetate (rho = 0.63); methional and methionol (rho = 0.60), but also with these two sulphur compounds and 2-phenylethan-1-ol with rhos 0.51 and 0.52 respectively. Additionally, a strong link

was found between 2-methyl-3-furanthiol and beta-damascenone l (rho = 0.58).

However, total understanding of the interactions is hard due to the multiple links between single pairs. Therefore, we further visualized both results with network layout based on the total edges originating from the raw values correlation analysis (Figures 7 and 8). Visualizing the links it is obvious that sub-groups exists, one obvious subcluster is Group 2 already found in Figure 5 composed of volatile phenols decanoic acid, butan-1-ol and acetaldehyde (Figure 7). Interestingly, it is composed of stronger odorants which are indeterminate for vineyards or fermentation location importances This group is mostly found to link with the fermentation location with 4-EG and butan-1-ol that are found important there although they appear in concentrations of limited sensory impact. The key central compound seem to be methional, methionol, beta-damascenone, 3SHA, hexan-1-ol, hexyl acetate and 2-phenylacetaldehyde group. Of these central compounds, only 3SHA is not found to be associated with fermentation location or vineyard differentiation based on the random forest feature selection. The location dependent (Z)-3-hexenol,the indeterminate ethyl propanoate and ethyl 2-methylbutanoate seem to associate with 3-sulfanylhexan-1-ol (3SH) which is found to be important distinguisher for both vineyards and fermentation location (Figure 7).

For the network edges based on relative abundances, many subgroups of aroma compounds seem exist. While few compounds are considered to be central for all the compositions: methional - methional interactions as well as the hexan-1-ol, hexyl acetate and 2-phenylacetaldehyde interaction. Interestingly 2-phenylacetaldehyde, beta-damascenone and 3SHA are linked situated close without apparent significant correlations together. Furthermore, 3SHA and hexyl acetate share close space, as do the 3SH and (Z)-3-hexenol (Figure 7).

Interestingly, while no real apparent sensory influence, three phenolic compounds: 4-ethyl guaiacol, 2-methoxyphenol and 2-methylphenol are found important in distinguishing the fermentation location, where they are found in higher amounts in the winery samples than in the aseptic ferments.

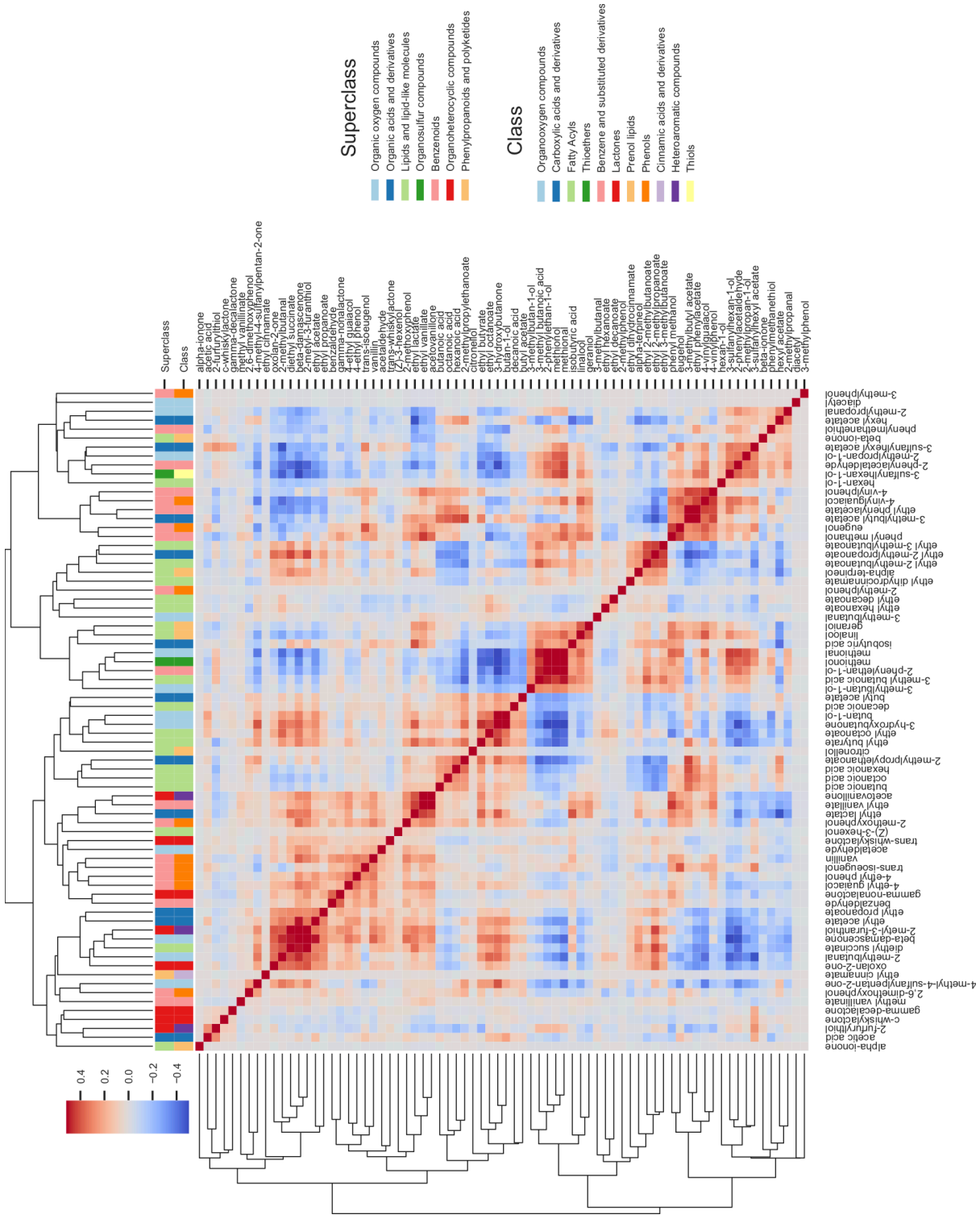


Figure 6. Clustered heatmap of chemical similarities based on OAV relative abundance of each wine and Bayesian compositional correlation analysis. Clustering using 'complete' method was done on the median estimates transformed to distances. Superclass and class correspond to classyfire (Djoumbou Feunang et al., 2016) taxonomy annotations.

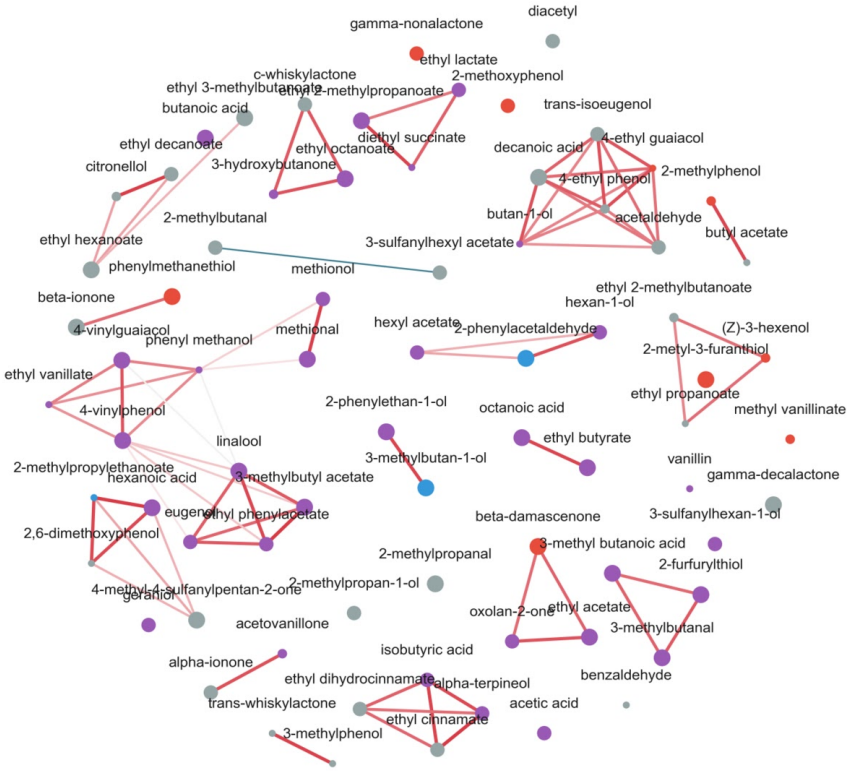


Figure 7. The network map of the co-occurring features' relative concentration pairwise correlations based on Bayesian compositional correlation analysis. Only significant edges with higher than rho ∈]-0.15,0,15[are shown absolute distance based kamada kawai layout. The node colors correspond to feature selected importance - red: found important for location classification; blue: found important for vineyard classification; purple: found important for both location and vineyard classifications; grey: not found important. The node sizes correspond to the mean odour activity values found in the wines.

Interestingly, while no real apparent sensory influence, three phenolic compounds: 4-ethyl guaiacol, 2-methoxyphenol and 2-methylphenol are found important in distinguishing the fermentation location, where they are found in higher amounts in the winery samples than in the aseptic ferments.

For the OAV abundance based compounds interactions, out of the 117 edges, we only visualize the significant interactions where one of the edge has OAV higher than 1, resulting to 50 edges (Figure 8). Four central compounds emerge: methionol, 2-phenylacetaldehyde, beta-damascenone and 2-methyl-3-furanthiol. These interactions seem to be further divided to behaviour where methionol and 2-phenylacetaldehyde relate negatively to most compounds except the alcohols: 2-phenylethan-1-ol, 3-sulfanylhexan-1-ol and 2-methylpropan-1-ol as well as methional. Whereas, beta-damascenone and 2-methyl-3-furanthiol together seem to have a more positive contribution to aroma compounds with the exception of 2-phenylacetaldehyde. These four compounds suggests that they interact with other aroma compounds either enhancing the odor activity in mixture (beta-damascenone and 2-methyl-3-furanthiol), or suppressing it (methionol, 2-phenylacetaldehyde).

Additionally it is interesting that linalool not positively correlate with the other main measured monoterpene geraniol but also with 4-VG and methional. Interestingly 4-VG seems to be antagonist with the thiol 4-methyl-4-sulfanylpentan-2-one (4MSP). 4MSP was negatively linking with the methional.

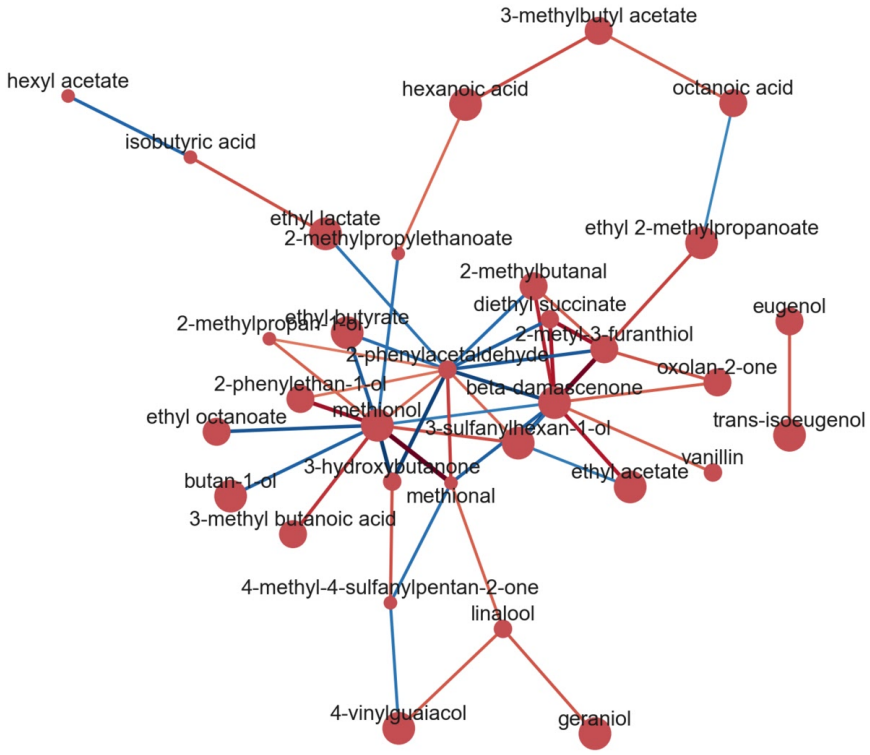


Figure 8. The network map of pairwise correlations of OAV transformed relative abundance. Only significant edges with higher than rho ∈]-0.15,0,15[and where one node has higher mean OAV value than 1 are shown. The node sizes correspond to the mean odour activity values found in the wines.

3.4 Linking chemistry and sensory

After the non-linear feature selection, we chose to investigate all the chemical compounds that belonged to a floral group in the FooDB. Four compounds, phenyl methanol, beta-damascenone, geraniol and linalool, were found. The concentration values were found to increase when the fermentation happened in the corresponding wineries (Figure 9). Two of these beta-damascenone and linalool were observed in values higher than their reported thresholds. Furthermore the thresholds for beta-damascenone as suggested by Pineau (Pineau et al., 2007) are also within the range of the sensory relevance of the compound. Furthermore geraniol is probably within the perception threshold of many people as at 20 µg/L 50 % of people are although to be able to smell it. Yet the phenyl methanol threshold, might be also influencing as well even though it is reported at 2 mg/L in a 10 % ethanol-water pH adjusted to 3.2 mixture (Culleré et al., 2004).

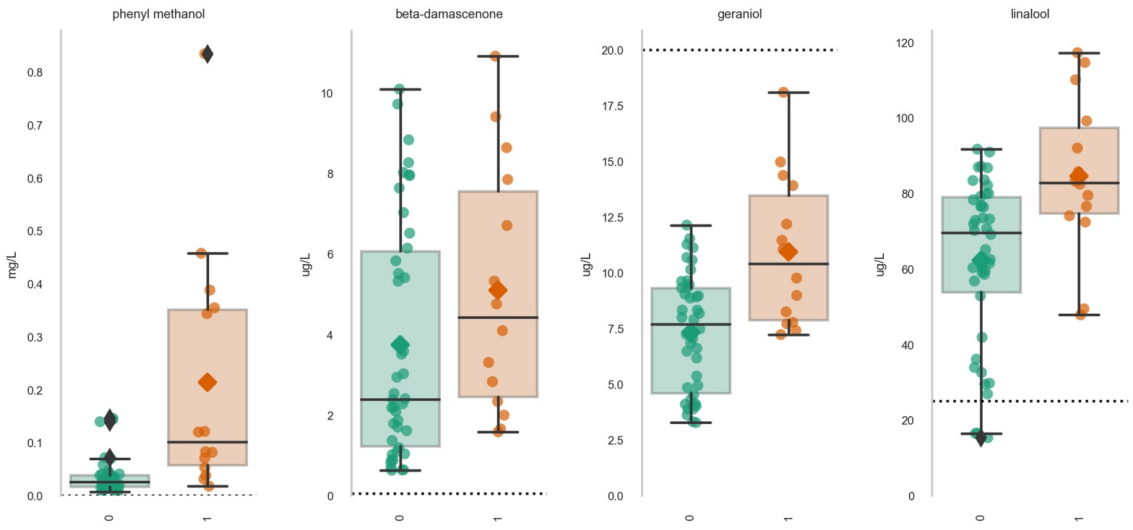


Figure 9. Boxplots of feature selected chemical compounds (phenyl methanol, beta-damascenone, geraniol and linalool) that had floral descriptor linked with their corresponding odour thresholds. For phenyl methanol the threshold, 2 mg/L (Culleré et al., 2004), has been changed to zero for visualization purposes. Fermentation location: 0 = vineyard, 1 = winery.

4 Discussion

Influence of wine microbiome on wine's sensory properties is crucial in wine making. However, studying these complex microbial processes and relating them to chemicals found in the final product is not straight forward with traditional analysis methods. In this study by using novel computational approaches we have revealed that sensorial and chemical compositions of wine is heavily influenced by the location of the alcoholic fermentation, vineyard and the year of harvest. We also show that there is a huge interaction between volatile compounds. These interactions can be translated into sensory relevance by using our novel approach that is based on odour activity abundance and compositionality.

Our results support earlier findings that there is enough Saccharomyces present in the vineyard to be able to carry out fermentations of the wines to dryness (Sirén et al. 2019b). Further, the winery fermented wines exhibited more fermentation problems than wines fermented in research winemaking scale conditions. This might be because of the interactions between the microbial communities that are derived from the vineyard and the winery. In the research winemaking scale conditions only the vineyard microbes are present and therefore there different population dynamics allowing fermentation completion (Boynton and Greig 2016). Additionally, if already white winemaking is highly different in respect to microbial communities, chances are that these differences are pronounced in red wines since they are even more exposed to the microbial community of the winery due to longer processing times. These findings however need to be confirmed in the future.

4.1 Network analyses

In order to understand how the sensory relevant measured volatile compounds influence each other through complex chemical interactions, we applied Bayesian approach to compositionality, and used derived correlations to infer a holistic view on the chemical composition. Based on such approach, the network analysis provided multiple interesting

pairwise interactions, which most of are already known through more focused studies. Examples such as oxidation and amino acid composition relationship between methional, methionol and 2-phenylethanol are established (Escudero et al., 2002; Hernández-Orte et al., 2005). For the two studied pairwise relationship approaches, the scaled concentration of compounds and the transformed OAV abundances, the strongest link was observed with the same pair of compounds: 3-methylbutyl and ethyl phenylacetate, which could also relate to classical winemaking conditions such as nitrogen composition (Antalick et al., 2014).

Another observed big cluster of compounds with the volatile ethyl phenols and fatty acids, most likely relates to microbial contamination. These compounds have been linked previously to microbial contamination issues such as *Brettanomyces* sp. (Joseph et al., 2015) and other "phenolic yeasts" (Shinohara, Kubodera, & Yanagida, 2000). Interestingly, these compounds were found in higher concentrations in wines fermented in wineries, allowing the winery microbial community to interact. An additional link between non-conventional microbial activity might relate to the positive correlation found with citronellol and ethyl decanoate through the various pathways (Spaepen et al., 1978; Vaudano et al., 2004).

Albeit in these observations the concentrations of the compounds were not found in sensory relevant levels. Although, we found in addition to these compounds other markers that suggest microbial differences between the two fermentations location effects. Such markers, ethyl lactate and diethyl succinate implicate a lactic acid bacteria contamination with the potential for malolactic fermentation. Only wines made in the wine estates were observed to undergo malolactic fermentation through changes in malic-lactic acid ratio, which might suggest that lactic acid bacteria did enter the vinification during processing. Therefore we underline that further investigations to the vineyard and winery microbial community interactions should be done.

The volatile thiols, 3SH and 3SHA are interesting molecules because of their strong odour activity values. They are also perceived positive odours, and especially 3SH has been suggested to be important part in Riesling typicality (Schüttler et al., 2015). We observed that 3SH has multiple origins in both winemaking and vineyard content, all of which are well-documented (Dubourdieu et al., 2006; Tominaga et al., 2000). We also

observed close interactions with both the c_6-alcohols and the c_6-alcohol acetates with the 3SHA. In general, we find the acetate of the 3SH is behaving similarly as multiple other acetates thus confirming previous observations (Swiegers et al., 2006). Thiols were observed in multiple occasions to be antagonists to oxidation related compounds, such as aldehydes. Such effect was found for two sensory relevant compounds: 2-methylbutanal, that negatively correlated with 3SHA, and methional, that correlated negatively with 4MSP and 3SH. The relationship between thiols and oxidation risks is known in general (Nikolantonaki et al., 2014) and furthermore methional has been found to suppress 3SH in sensory interactions (Coetzee et al., 2015), thus confirming our observations. In our study these inhibiting effects were found to be crucial even when the levels of 2-methylbutanal is under sensory thresholds.

Therefore regarding the perceivable odor, two effects exists: chemical suppression that we have observed in this study but also the suppressing effect directly tied to the sensory perception (Pineau et al., 2007). Interestingly, the same chemical compounds might impact both outcomes. As an example, beta-damascenone, a plant derived c_{13}-norisoprenoid that is a key wine odorant as reviewed extensively (Sefton et al., 2011). Additionally it has been shown interacting with other odorants enhancing fruit aroma perceptions (Pineau et al., 2007), whereas we suggest that it also influences the sensory sphere more directly influencing the other compounds through the chemical composition

Another compound that we found to behave in a similar fashion to beta-damascenone is 2-methyl-3-furanthiol (Figure 8). This compound has been evidenced previously to act as meat aroma enhancer (Meng et al., 2017). While both of these compounds acted positively on aroma, we also found two compounds with negative influence: methionol and 2-phenylacetaldehyde. Previous studies have shown that these compounds can indeed suppress aroma (Culleré et al., 2007; Silva Ferreira et al., 2003). All above examples demonstrate that our compositional OAV abundance approach with using large amount of data allows generating hypotheses of potential sensory perceptions from the data directly. This approach offers promising directions to improve the understanding of the chemical-sensory interface of complex mixtures such as wine.

Multiple issues however needs to be investigated before this approach can be used in an efficient manner. In general, the relative abundance offers promise while avoiding the

common pitfalls of using OAV to explaining odour intensity (Grosch, 2001). The OAVs were originally not even recommended for intensity measures but rather compositionality (Guadagni et al., 1966). Additional issue with ODT remains a critical issue for both compositionality and intensity based approaches. While the actual Bayesian approach seems to be relatively robust in investigating the compositional changes, the follow-up downstream analyses can be seriously misleading, if the OAV-values are not accurately established. When comparing the raw feature transformation and the OAV abundance approaches, it is evident that while multiple pairwise relationships are found with both methods, information differences can be huge, and from which multiple found correlations might not be transformed into sensory relevance. While linking sensory perceptions with chemical perceptions is possible, usually the values need to be transformed to their corresponding odor activity values. Thus further work is needed in order to carry out proper investigations on the whole matrix.

Furthermore by using novel machine learning approaches and ontology-based inter-study associations, easy explanations can surface such as deriving the aroma precursors or the odour type, as shown with the chemical explanation of floral, honey descriptors associated with wineries (Figure 9). Interestingly it seems to validate that the compound co-behaviour is in many cases closely linked to the precursor similarity as well as structural similarity. In our study it was found that in wine only few generalists (Dunkel et al., 2014) play a key role when measured as relevant abundances.

4.2 Microbial community interactions relate to sensory perceptions

The findings of this study lead to better understanding how to control the native microbial populations when aspiring after specific wine styles. We hypothesise that when considering true microbial regionality in wine, the interaction between cellar flora and vineyard flora needs to be taken into account as it can cause large variation between wines, purposeful or not. We found that the vineyards are highly differentiated in their chemical compositions. This tends to be less obvious with a more robust method such as descriptive sensory analysis. Although vineyards vary in their microbial composition they also differ between their precursor composition derived from the grape, which might cause the observed differences. Yet, the separation of the vineyards became less obvious

with the samples from wineries which could suggest that the winery microbiome has a crucial role in the final aroma profile. Furthermore both the year and winery seemed to carry more impact than the originating vineyard when measured in terms of introduced variation. Where year effect was expected, the winery effect was not expected to be this strong.

Raw data and feature selection of relevant sensory descriptors for fermentation location microbial population (vineyard or cellar and vineyard) allowed to derive information about sensory differentiation of wines. In general wines fermented by the cellar and vineyard microbial populations were perceived more fruit-driven, exhibited honey and floral odors with higher intensities whereas the wines fermented with solely vineyard population were more acidity-driven and had less obvious odours. This was validated with targeted aroma analysis. Four chemical compounds (linalool, geraniol, 2-methoxyphenyl, 2-phenylacetaldehyde) that are well-known to exhibit floral odour were found important by using feature selection of the FooDB ontologies in our analysis, and further found to be in higher concentrations in the winery wines validating the feature based approach.

Here we have linked the chemical compounds in two ways. First, by investigating their co-occurrence relationships with having same weight on all the features. This has allowed us to associate the compounds together in order to derive their behaviour similarity. Secondly, we investigated the importance of the compounds to wine aroma composition. This we have done by centering the focusing on the odour activity, which we believe is composed not only of compound odour intensities per se but also of their co-occurrence and relative abundance. Further studies could investigate the complexity and interactions of these by looking at diversity and richness, evenness estimates inside the sample but also by comparing the diversities across samples. Such approaches are commonly applied in biology and ecology studies (Chao et al., 2004; Legendre and Legendre, 2012; Shannon, 1948; Simpson, 1949).

While we observe that the physico-chemical parameters between carboys and winery fermentors might play a role in differentiating the wines, they cannot explain all observations, such as the appearance of lactic acid or volatile ethyl phenols. In order to establish a stronger metabolic link between the precursor and the observed compounds,

the two Bayesian approaches used in this study could be further investigated with a third approach, where the links between actual concentrations of the compounds are determined. Such work is currently ongoing. However, it can be concluded from our studies that a microbial population effect between the sensory and chemical realm does exist.

5 Conclusion

This study shed light upon the role of the vineyard and winery microbial communities in aroma compound production and on how the two communities interact. This study is the first to show that such relationships do exist in holistic environment, and not just *in vitro*. By investigating the interaction of aroma compounds, the present findings underline that proper chemical analysis equipped with data analysis can generate but also validate new hypotheses that link sensory-chemical interface. Further holistic studies are still needed to understand how compounds interact with each other,, which might help us also to establish more solid relationships between chemical composition and sensory descriptions. The hypotheses generated by such studies can then be further investigated with more targeted studies.

Non-targeted multi-omics approaches *in situ* could be useful to derive further biological explanations and biological relationships. By using tools such as shotgun sequencing and non-targeted metabolomics, functional analysis profiling of the microbial community activity patterns could elucidate the actual biological pathways associating to the formation of specific sensory relevant aroma compounds. Such novel findings could afterwards be verified with targeted approaches.

Acknowledgments

The research was funded by Horizon 2020 Programme of the European Commission within the Marie Skłodowska-Curie Innovative Training Network “MicroWine” (grant number 643063). The authors would like to thank Sandra Klink in assisting the sensory analysis set up and all the sensory panelists during 2016 and 2017, both harvest teams in 2015 and 2016, Laura Culleré and whole team at LAEE for doing the GC-MS analyses in

Zaragoza, Pascal Wegmann-Herr, Sascha Wolz, Benedikt Grein and Florian Schraut for accommodating the winemaking facilities. Hannah Iberer for assisting during the fermentations in 2016. Sarah Mak and Tiia Kittilä for assistance and proofreading the manuscript draft. All the collaborating wineries and winemakers in Pfalz.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.

Antalick, G., Perello, M.-C., and de Revel, G. (2014). Esters in Wines: New Insight through the Establishment of a Database of French Wines. *Am. J. Enol. Vitic.*, ajev.2014.13133.

Hagberg, A. A., Schult, D. A. and Swart, P. J. (2008) Exploring network structure, dynamics, and function using NetworkX. in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, (Eds). Varoquaux G., Vaught T., and Millman J. 11–15.

Bauer, A., Wolz, S., Schormann, A., and Fischer, U. (2011). 'Authentication of Different Terroirs of German Riesling Applying Sensory and Flavor Analysis', in *Progress in Authentication of Food and Wine* ACS Symposium Series. (American Chemical Society), 131–149.

Belda, I., Ruiz, J., Alastruey-Izquierdo, A., Navascués, E., Marquina, D., and Santos, A. (2016). Unraveling the Enzymatic Basis of Wine 'Flavorome': A Phylo-Functional Study of Wine Related Yeast Species. *Front. Microbiol.* 7, 243.

Belda, I., Ruiz, J., Esteban-Fernández, A., Navascués, E., Marquina, D., Santos, A., et al. (2017a). Microbial Contribution to Wine Aroma and Its Intended Use for Wine Quality Improvement. *Molecules* 22, 189.

Belda, I., Zarraonaindia, I., Perisin, M., Palacios, A., and Acedo, A. (2017b). From Vineyard Soil to Wine Fermentation: Microbiome Approximations to Explain the 'terroir' Concept. *Front. Microbiol.* 8, 8966.

Bisson, L. F. (1999). Stuck and Sluggish Fermentations. *Am. J. Enol. Vitic.* 50, 107–119.

Bodenhofer, U., Kothmeier, A., and Hochreiter, S. (2011). APCluster: an R package for affinity propagation clustering. *Bioinformatics* 27, 2463–2464.

Bokulich, N. A., Collins, T. S., Masarweh, C., Allen, G., Heymann, H., Ebeler, S. E., et al. (2016). Associations among Wine Grape Microbiome, Metabolome, and Fermentation Behavior Suggest Microbial Contribution to Regional Wine Characteristics. *MBio* 7. doi:10.1128/mBio.00631-16.

Bokulich, N. A., Ohta, M., Richardson, P. M., and Mills, D. A. (2013). Monitoring Seasonal Changes in Winery-Resident Microbiota. *PLoS One* 8, e66437.

Bokulich, N. A., Thorngate, J. H., Richardson, P. M., and Mills, D. A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc. Natl. Acad. Sci. U. S. A.* 111, E139–48.

Borneman, A. R., Forgan, A. H., Kolouchova, R., Fraser, J. A., and Schmidt, S. A. (2016). Whole Genome Comparison Reveals High Levels of Inbreeding and Strain Redundancy Across the Spectrum of Commercial Wine Strains of Saccharomyces cerevisiae. *G3* 6, 957–971.

Boynton, P. J., and Greig, D. (2016). Species richness influences wine ecosystem function through a dominant species. *Fungal Ecol.* 22, 61–72.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *J. Math. Sociol.* 25, 163–177.

Bueno, M., Zapata, J., and Ferreira, V. (2014). Simultaneous determination of free and bonded forms of odor-active carbonyls in wine using a headspace solid phase microextraction strategy. *J. Chromatogr. A* 1369, 33–42.

Campo, E., Ferreira, V., and Escudero, A. (2005). Prediction of the wine sensory properties related to grape variety from dynamic-headspace gas chromatography− olfactometry data. *Journal of Agricultural*. Available at: https://pubs.acs.org/doi/abs/10.1021/jf047870a.

Chao, A., Chazdon, R. L., Colwell, R. K., and Shen, T.-J. (2004). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.* 8, 148–159.

Chastrette, M., and Rallet, E. (1998). Structure--minty odour relationships: suggestion of an interaction pattern. *Flavour Fragr. J.* 13, 5–18.

Ciani, M., Capece, A., Comitini, F., Canonico, L., Siesto, G., and Romano, P. (2016). Yeast Interactions in Inoculated Wine Fermentation. *Front. Microbiol.* 7, 555.

Ciani, M., Mannazzu, I., Marinangeli, P., Clementi, F., and Martini, A. (2004). Contribution of winery-resident Saccharomyces cerevisiae strains to spontaneous grape must fermentation. *Antonie Van Leeuwenhoek* 85, 159–164.

Coetzee, C., Brand, J., Emerton, G., Jacobson, D., Silva Ferreira, A. C., and du Toit, W. J. (2015). Sensory interaction between 3-mercaptohexan-1-ol, 3-isobutyl-2-methoxypyrazine and oxidation-related compounds: Sauvignon Blanc sensory interaction studies. *Aust. J. Grape Wine Res.* 21, 179–188.

Culleré, L., Cacho, J., and Ferreira, V. (2007). An assessment of the role played by some oxidation-related aldehydes in wine aroma. *J. Agric. Food Chem.* 55, 876–881.

Culleré, L., Escudero, A., Cacho, J., and Ferreira, V. (2004). Gas chromatography-olfactometry and chemical quantitative study of the aroma of six premium quality spanish aged red wines. *J. Agric. Food Chem.* 52, 1653–1660.

Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., et al. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminform.* 8, 61.

Dubourdieu, D., Tominaga, T., Masneuf, I., des Gachons, C. P., and Murat, M. L. (2006). The Role of Yeasts in Grape Flavor Development during Fermentation: The Example of Sauvignon blanc. *Am. J. Enol. Vitic.* 57, 81–88.

Dunkel, A., Steinhaus, M., Kotthoff, M., Nowak, B., Krautwurst, D., Schieberle, P., et al. (2014). Nature's Chemical Signatures in Human Olfaction: A Foodborne Perspective for Future Biotechnology. *Angew. Chem. Int. Ed.* 53, 7124–7143.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C. (2003). Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* 35, 279–300.

Escudero, A., Asensio, E., Cacho, J., and Ferreira, V. (2002). Sensory and chemical changes of young white wines stored under oxygen. An assessment of the role played by aldehydes and some other important odorants. *Food Chem.* 77(3), 325-331.

Ferreira, V., Sáenz-Navajas, M.-P., Campo, E., Herrero, P., de la Fuente, A., and Fernández-Zurbano, P. (2016). Sensory interactions between six common aroma vectors explain four main red wine aroma nuances. *Food Chem.* 199, 447–456.

Fischer, U. (2007). 'Wine Aroma', in *Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability*, ed. R. G. Berger (Springer Berlin Heidelberg), 241–267.

Fischer, U., Berger, R. G., Hakansson, A., and Noble, A. C. (1996). The impact of dealcoholization on the flavour of wine - relating concentration of aroma compounds to sensory data using PLS analysis. *Flavour Science. Recent Developments* 197, 335–338.

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry* 40, 35–41.

Frey, B. J., and Dueck, D. (2007). Clustering by passing messages between data points. *Science* 315, 972–976.

Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., et al. (2016). Domestication and Divergence of Saccharomyces cerevisiae Beer Yeasts. *Cell* 166, 1397–1410.e16.

Grosch, W. (2001). Evaluation of the key odorants of foods by dilution experiments, aroma models and omission. *Chem. Senses* 26, 533–545.

Guadagni, D. G., Buttery, R. G., and Harris, J. (1966). Odour intensities of hop oil components. *J. Sci. Food Agric.* 17, 142–144.

Heller, S. R., McNaught, A., Pletnev, I., Stein, S., and Tchekhovskoi, D. (2015). InChI, the IUPAC International Chemical Identifier. *J. Cheminform.* 7, 23.

Hernández-Orte, P., Ibarz, M. J., Cacho, J., and Ferreira, V. (2005). Effect of the addition of ammonium and amino acids to musts of Airen variety on aromatic composition and sensory properties of the obtained wine. *Food Chem.* 89, 163–174.

Hopfer, H., Nelson, J., Collins, T. S., Heymann, H., and Ebeler, S. E. (2015). The combined impact of vineyard origin and processing winery on the elemental profile of red wines. *Food Chem.* 172, 486–496.

Hothorn, T., Hornik, K., van de Wiel, M. A., and Zeileis, A. (2006). A Lego System for Conditional Inference. *Am. Stat.* 60, 257–263.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9, 90–95.

Oksanen J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2018). *vegan: Community Ecology Package*.

Joseph, L. J. C. M., Albino, E. A., Ebeler, S. E., and Bisson, L. F. (2015). Brettanomyces bruxellensis Aroma Active Compounds Determined by SPME GC-MS Olfactory Analysis. *Am. J. Enol. Vitic.*, ajev. 2015.14073.

Juan, F. S., Cacho, J., Ferreira, V., and Escudero, A. (2013). Flavour Science: Chapter 21. Differences in Chemical Composition of Aroma among Red Wines of Different Price *J. Agric. Food Chem.* 2012, 60, 20, 5045-5056

Knight, S. J., Klaere, S., Morrison-Whittle, P., and Goddard, M. R. (2018). Fungal diversity during fermentation correlates with thiol concentration in wine: Fungal diversity correlates with wine thiol levels. *Aust. J. Grape Wine Res.* 24, 105–112.

Knight, S., Klaere, S., Fedrizzi, B., and Goddard, M. R. (2015). Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir. *Sci. Rep.* 5, 14233.

Kraus, J. K., Reed, G., and Villettaz, J.-C. (1984). Levures sèches actives de vinification. 2e partie et fin: utilisation et évaluation. *OENO One* 18, 1.

Kursa, M. B., Rudnicki, W. R., and Others (2010). Feature selection with the Boruta package. *J. Stat. Softw.* 36, 1–13.

Lawless, H. T., and Heymann, H. (2010). *Sensory Evaluation of Food: Principles and Practices*. Springer Science & Business Media.

Lee, S.-J., and Noble, A. C. (2003). Characterization of odor-active compounds in Californian chardonnay wines using GC-olfactometry and GC-mass spectrometry. *J. Agric. Food Chem.* 51, 8036–8044.

Lee, S.-J., and Noble, A. C. (2006). Use of Partial Least Squares Regression and Multidimensional Scaling on Aroma Models of California Chardonnay Wines. *Am. J. Enol. Vitic.* 57, 363–370.

Legendre, P., and Legendre, L. (2012). ‘Chapter 1 - Complex ecological data sets’, in *Developments in Environmental Modelling*, eds. P. Legendre and L. Legendre (Elsevier), 1–57.

Legendre, P., Oksanen, J., and ter Braak, C. J. F. (2011). Testing the significance of canonical axes in redundancy analysis. *Methods Ecol. Evol.* 2, 269–277.

Le, S., and Husson, F. (2008). Sensominer: A Package For Sensory Data Analysis. *J. Sens. Stud.* 23, 14–25.

López, R., Aznar, M., Cacho, J., and Ferreira, V. (2002). Determination of minor and trace volatile compounds in wine by solid-phase extraction and gas chromatography with mass spectrometric detection. *J. Chromatogr. A* 966, 167–177.

Louw, L., Malherbe, S., Naes, T., Lambrechts, M., van Rensburg, P., and Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Qual. Prefer.* 30, 192–201.

Mak, S. S. T., Sirén, K., Carøe, C., Gopalakrishnan, S., Ellegaard, M. R., Klincke, F., Hansen, J. H., Fische, U. and Gilbert M. T. P. (2019 under review) Comparison of DNA extraction methods for use in fungal diversity analyses of Riesling during alcoholic fermentation.

Malherbe, S., Bauer, F. F., and Du Toit, M. (2007). Understanding Problem Fermentations - A Review. *S. Afr. J. Enol. Vitic.* 28(2), 169-186.

Mateo-Vivaracho, L., Zapata, J., Cacho, J., and Ferreira, V. (2010). Analysis, occurrence, and potential sensory significance of five polyfunctional mercaptans in white wines. *J. Agric. Food Chem.* 58, 10184–10194.

McKinney, W. (2010). Data Structures for Statistical Computing in Python. in *Proceedings of the 9th Python in Science Conference*, 51–56.

Meng, Q., Kitagawa, R., Imamura, M., Katayama, H., Obata, A., and Sugawara, E. (2017). Contribution of 2-methyl-3-furanthiol to the cooked meat-like aroma of fermented soy sauce. *Biosci. Biotechnol. Biochem.* 81, 168–172.

Mezzasalma, V., Sandionigi, A., Guzzetti, L., Galimberti, A., Grando, M. S., Tardaguila, J., et al. (2018). Geographical and Cultivar Features Differentiate Grape Microbiota in Northern Italy and Spain Vineyards. *Front. Microbiol.* 9, 946.

Morrison-Whittle, P., and Goddard, M. R. (2018). From vineyard to winery: a source map of microbial diversity driving wine fermentation. *Environ. Microbiol.* 20, 75–84.

Naes, T., Brockhoff, P. B., and Tomic, O. (2010). *Statistics for Sensory and Consumer Science*. Chichester, UK: John Wiley & Sons, Ltd.

Nikolantonaki, M., Magiatis, P., and Waterhouse, A. L. (2014). Measuring protection of aromatic wine thiols from oxidation by competitive reactions vs wine preservatives with ortho-quinones. *Food Chem.* 163, 61–67.

Noble, A. C., and Ebeler, S. E. (2002). Use of multivariate statistics in understanding wine flavor. *Food Rev. Int.* 18, 1–20.

Nozaki, Y., and Nakamoto, T. (2018). Predictive modeling for odor character of a chemical using machine learning combined with natural language processing. *PLoS One* 13, e0198475.

Ortega, C., López, R., Cacho, J., and Ferreira, V. (2001). Fast analysis of important wine volatile compounds: Development and validation of a new method based on gas chromatographic–flame ionisation detection analysis of dichloromethane microextracts. *J. Chromatogr. A* 923, 205–214.

Padilla, B., Gil, J. V., and Manzanares, P. (2016). Past and Future of Non-Saccharomyces Yeasts: From Spoilage Microorganisms to Biotechnological Tools for Improving Wine Aroma Complexity. *Front. Microbiol.* 7, 411.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pineau, B., Barbe, J.-C., Van Leeuwen, C., and Dubourdieu, D. (2007). Which impact for beta-damascenone on red wines aroma? *J. Agric. Food Chem.* 55, 4103–4108.

Pineau, B., Barbe, J.-C., Van Leeuwen, C., and Dubourdieu, D. (2009). Examples of Perceptive Interactions Involved in Specific 'Red-' and 'Black-berry' Aromas in Red Wines. *J. Agric. Food Chem.* 57, 3702–3708.

Poivet, E., Tahirova, N., Peterlin, Z., Xu, L., Zou, D.-J., Acree, T., et al. (2018). Functional odor classification through a medicinal chemistry approach. *Sci Adv* 4, eaao6086.

Quan, A. S., and Eisen, M. B. (2018). The ecology of the Drosophila-yeast mutualism in wineries. *PLoS One* 13, e0196440.

Rapp, A., and Mandery, H. (1986). Wine aroma. *Experientia* 42, 873–884.

Reimann, C., Filzmoser, P., Hron, K., Kynčlová, P., and Garrett, R. G. (2017). A new method for correlation analysis of compositional (environmental) data--a worked example. *Sci. Total Environ.* 607, 965–971.

Ribereau Gayon, P. (1999). Observations related to the causes and consequences of stuck fermentation in vinification [aeration of must, control of temperature, nutrient starvation]. *International Journal of Vine and Wine Sciences* 33, 39–48.

Rossouw, D., Bagheri, B., Setati, M. E., and Bauer, F. F. (2015). Co-Flocculation of Yeast Species, a New Mechanism to Govern Population Dynamics in Microbial Ecosystems. *PLoS One* 10, e0136249.

Sadoudi, M., Tourdot-Maréchal, R., Rousseaux, S., Steyer, D., Gallardo-Chacón, J.-J., Ballester, J., et al. (2012). Yeast-yeast interactions revealed by aromatic profile analysis of Sauvignon Blanc wine fermented by single or co-culture of non-Saccharomyces and Saccharomyces yeasts. *Food Microbiol.* 32, 243–253.

Sáenz-Navajas, M.-P., Arias, I., Ferrero-Del-Teso, S., Fernández-Zurbano, P., Escudero, A., and Ferreira, V. (2018). Chemo-sensory approach for the identification of chemical compounds driving green character in red wines. *Food Res. Int.* 109, 138–148.

Sáenz-Navajas, M.-P., Campo, E., Culleré, L., Fernández-Zurbano, P., Valentin, D., and Ferreira, V. (2010). Effects of the nonvolatile matrix on the aroma perception of wine. *J. Agric. Food Chem.* 58, 5574–5585.

San-Juan, F., Ferreira, V., Cacho, J., and Escudero, A. (2011). Quality and aromatic sensory descriptors (mainly fresh and dry fruit character) of Spanish red wines can be predicted from their aroma-active chemical composition. *J. Agric. Food Chem.* 59, 7916–7924.

Schüttler, A., Friedel, M., Jung, R., Rauhut, D., and Darriet, P. (2015). Characterizing aromatic typicality of Riesling wines: merging volatile compositional and sensory aspects. *Food Res. Int.* 69, 26–37.

Schwager, E. (2015). banocc: Bayesian ANalysis Of Compositional Covariance. R package version 1.2.0

Schwager, E., Mallick, H., Ventz, S., and Huttenhower, C. (2017). A Bayesian method for detecting pairwise associations in compositional data. *PLoS Comput. Biol.* 13, e1005852.

Sefton, M. A., Skouroumounis, G. K., Elsey, G. M., and Taylor, D. K. (2011). Occurrence, sensory impact, formation, and fate of damascenone in grapes, wines, and other foods and beverages. *J. Agric. Food Chem.* 59, 9717–9746.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal* 27, 379–423.

Sherman, E., Harbertson, J. F., Greenwood, D. R., Villas-Bôas, S. G., Fiehn, O., and Heymann, H. (2018). Reference samples guide variable selection for correlation of wine sensory and volatile profiling data. *Food Chem.* 267, 344–354.

Shinohara, T., Kubodera, S., and Yanagida, F. (2000). Distribution of Phenolic Yeasts and Production of Phenolic Off-Flavors in Wine Fermentation. *J. Biosci. Bioeng.* 90, 90–97.

Siebert, T. E., Barker, A., Pearson, W., Barter, S. R., de Barros Lopes, M. A., Darriet, P., et al. (2018). Volatile Compounds Related to 'Stone Fruit' Aroma Attributes in Viognier and Chardonnay Wines. *J. Agric. Food Chem.* 66, 2838–2850.

Silva Ferreira, A. C., Hogg, T., and Guedes de Pinho, P. (2003). Identification of key odorants related to the typical aroma of oxidation-spoiled white wines. *J. Agric. Food Chem.* 51, 1377–1381.

Simpson, E. H. (1949). Measurement of Diversity. *Nature* 163, 688.

Sirén, K., Mak, S. S. T., Fischer, U., Hansen, L. H., and Gilbert, M. T. P. (2019a). Multi-omics and potential applications in wine production. *Curr. Opin. Biotechnol.* 56, 172–178.

Sirén, K., Mak, S. S. T., Melkonian, C., Carøe, C., Swiegers, J. H., Molenaar, D., Fischer, U. and Gilbert, M. T. P. (2019b under review). Taxonomic and functional characterisation of the microbial community during spontaneous in vitro fermentation of Riesling must.

Spaepen, M., Van Oevelen, D., and Verachtert, H. (1978). Fatty Acids And Esters Produced During The Spontaneous Fermentation Of Lambic And Gueuze. *J. Inst. Brew.* 84, 278–282.

Stan Development Team (2018). RStan: the R interface to Stan. Available at: http://mc-stan.org/.

Stefanini, I., and Cavalieri, D. (2018). Metagenomic Approaches to Investigate the Contribution of the Vineyard Environment to the Quality of Wine Fermentation: Potentials and Difficulties. *Front. Microbiol.* 9, 991.

Swiegers, J. H., Willmott, R., Hill-Ling, A., Capone, D. L., Pardon, K. H., Elsey, G. M., et al. (2006). 'Modulation of volatile thiol and ester aromas by modified wine yeast', in *Developments in Food Science*, eds. W. L. P. Bredie and M. A. Petersen (Elsevier), 113–116.

Tempère, S., Marchal, A., Barbe, J.-C., Bely, M., Masneuf-Pomarede, I., Marullo, P., et al. (2018). The complexity of wine: clarifying the role of microorganisms. *Appl. Microbiol. Biotechnol.* 102, 3995–4007.

ter Braak, C. J. F. (1986). Canonical Correspondence Analysis: A New Eigenvector Technique for Multivariate Direct Gradient Analysis. *Ecology* 67, 1167–1179.

Tominaga, T., Baltenweck-Guyot, R., Des Gachons, C. P., and Dubourdieu, D. (2000). Contribution of Volatile Thiols to the Aromas of White Wines Made From Several Vitis vinifera Grape Varieties. *Am. J. Enol. Vitic.* 51, 178–181.

Varela, C. (2016). The impact of non-Saccharomyces yeasts in the production of alcoholic beverages. *Appl. Microbiol. Biotechnol.* 100, 9861–9874.

Varela, C., and Borneman, A. R. (2017). Yeasts found in vineyards and wineries. *Yeast* 34, 111–128.

Vaudano, E., Moruno, E. G., and Stefano, R. (2004). Modulation of Geraniol Metabolism During Alcohol Fermentation. *J. Inst. Brew.* 110, 213–219.

Wang, C., Mas, A., and Esteve-Zarzoso, B. (2016). The Interaction between Saccharomyces cerevisiae and Non-Saccharomyces Yeast during Alcoholic Fermentation Is Species and Strain Specific. *Front. Microbiol.* 7, 502.

Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lukauskas, S., Gemperline, D. C., et al. (2017). *mwaskom/seaborn: v0.8.1 (September 2017)*. doi:10.5281/zenodo.883859.

Winton, W., Ough, C. S., and Singleton, V. L. (1975). Relative Distinctiveness of Varietal Wines Estimated by the Ability of Trained Panelists to Name the Grape Variety Correctly. *Am. J. Enol. Vitic.* 26, 5–11.

Wishart, D. S. (2014). FooDB: the food database. FooDB version 1.0.

Wright, M., and Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software, Articles* 77, 1–17.

Wu, C., Liu, J., Zhao, P., Piringer, M., and Schauberger, G. (2016). Conversion of the chemical concentration of odorous mixtures into odour concentration and odour intensity: A comparison of methods. *Atmos. Environ.* 127, 283–292.

Xufre, A., Albergaria, H., Inácio, J., Spencer-Martins, I., and Gírio, F. (2006). Application of fluorescence in situ hybridisation (FISH) to the analysis of yeast population dynamics in winery and laboratory grape must fermentations. *Int. J. Food Microbiol.* 108, 376–384.

Table 1. Details of bottled samples. Mean values with reported standard deviations if notable.

Sample origin	Sample id	Year	Residual sugar (g/L)	Tartaric acid (g/L)	Lactic acid (g/L)	Malic acid (g/L)	Volatile acidity (g/L)
vineyard 1	wb 1 a-c	2015	1.8 ± 0.59	3,72	0,33	2,43	0,1
vineyard 2	wb 2 a-c	2015	6.4 ± 1.7	3,81	0,17	2,6	0,2
vineyard 3	wb 3 a-c	2015	7.4 ± 0.9	3,56	0,06	2,64	0,2
vineyard 4	wb 4 a-c	2015	6.6 ± 0.5	4,49	0,11	2,87	0,3
vineyard 5	wb 5 a-c	2015	5.8 ± 3.5	4,74	0	3,43	0,3
winery 1	wgt 1	2015	2,3	1,29	2,31	0,26	0,2
winery 2	wgt 2	2015	0,5	3,09	2,15	0,45	0,6
winery 3	wgt 3	2015	2,7	2,65	2,06	0,05	0,4
winery 4	wgt 4	2015	8	2,48	1,98	0,11	0,4
winery 5	wgt 5	2015	1	3,03	2,24	-0,02	0,3
vineyard 1	wb 1 a-c	2016	5.0 ± 0.3	3,7	0,1	2,1	0,1
vineyard 2	wb 2 a-c	2016	4.0 ± 1	2,7	0,2	2,1	0,1
vineyard 3	wb 3 a-c	2016	5.1 ± 0.5	3,3	0,3	2,8	0,2
vineyard 4	wb 4 a-c	2016	3.2 ± 0.5	3,9	0,1	3,4	0,2
vineyard 5	wb 5 a-c	2016	5.3 ± 0.6	4,1	0	3,6	0,2
vineyard 6	wb 6 a,c	2016	3.5 ± 0.2	3,6	0,1	2,1	0,2
vineyard 7	wb 7 a-c	2016	0.2 ± 0.3	3,2	0,4	1,9	0,1
winery 1	wgt 1	2016	2,9	2,4	0,6	1,6	0,1
winery 2	wgt 2	2016	8,6	2,2	0,7	1,8	0,2
winery 3	wgt 3	2016	5,5	3,5	2,2	1,6	0,4
winery 4	wgt 4	2016	5,9	2,6	0,1	3,1	0,2
winery 5	wgt 5	2016	2,4	2,6	0	4,1	0,3
winery 6	wgt 6 (t202, f77, f775)	2016	3.5 ± 2.7	3	0.4 ± 0.3	1.3 ± 1.2	0,1
winery 7	wgt 7 (t209, f30)	2016	1.5 ± 1.8	2,7	0.7 ± 0.6	2 ± 0.6	0,1
vineyard 1 yeast 1	17 a,b	2016	1.5 ± 1.0	3,7	0,1	2,58	0,1
vineyard 1 yeast 2	18 a,b	2016	0.3 ± 0.5	3,7	0,1	2,78	0,1
vineyard 1 40 mg SO_2	15 a-c	2016	5 ± 1.6	3,8	0,1	2,34	0,1

Table 2. Groups and subgroups of volatile compounds generated from the clustering of the Bayesian compositional correlation analysis.

Group 1	Group 2					
Subgroup1	Subgroup2	Subgroup1	Subgroup 2			
-	-	Cluster1	Cluster 2	Cluster 1	Cluster 2	Cluster 3
methionol methional linalool geraniol 3-methyl-butylacetate ethylphenyl-acetate 3-methyl-butan-1-ol 2-phenyl-ethan-1-ol phenyl methanol eugenol ethylvanillate methyl-vanillinate 4-vinyl-guaiacol 4-vinyl-phenol	3-methyl-butanoic acid 2-methyl-propan-1-ol 3-methyl-butanal 2phenyl-acetalde-hyde 2-methyl-propanal hexan-1-ol hexyl acetate 3-sulfanyl-hexan-1-ol 3-sulfanyl-hexyl acetate acetic acid diacetyl beta-ionone phenyl methanethiol 2-furfurylthiol	ethyl 3-methyl-butanoate ethyl 2-methyl-propanoate ethyl 2-methyl-butanoate ethylpropanoate ethylcinnamate ethyldihydro-cinnamate ethyllactate diethylsuccinate (Z)-3-hexenol isobutyric acid alpha-terpineol	3-methyl-phenol beta-damascenone 2-methyl-3-furanthiol ethyl acetate oxolan-2-one benzaldehyde gamma-nonalactone trans-whiskylactone vanillin	butan-1-ol decanoic acid acetaldehyde trans-isoeugenol 4-ethyl-guaiacol 4-ethyl-phenol	2-methoxy-phenol aceto-vanillone alphaionone ethylbutyrate butanoic acid octanoic acid	3-hydroxy-butanone 2-methylbutanal c-whiskylactone ethyl hexanoate ethyl octanoate ethyl decanoate 2-methyl-propylethanoate citronellol butyl acetate 2-methylphenol 2,6-dimethoxy-phenol hexanoic acid 4-methyl-4-sulfanylpentan-2-one

Supplementary information of Chapter 5

Supplementary figures

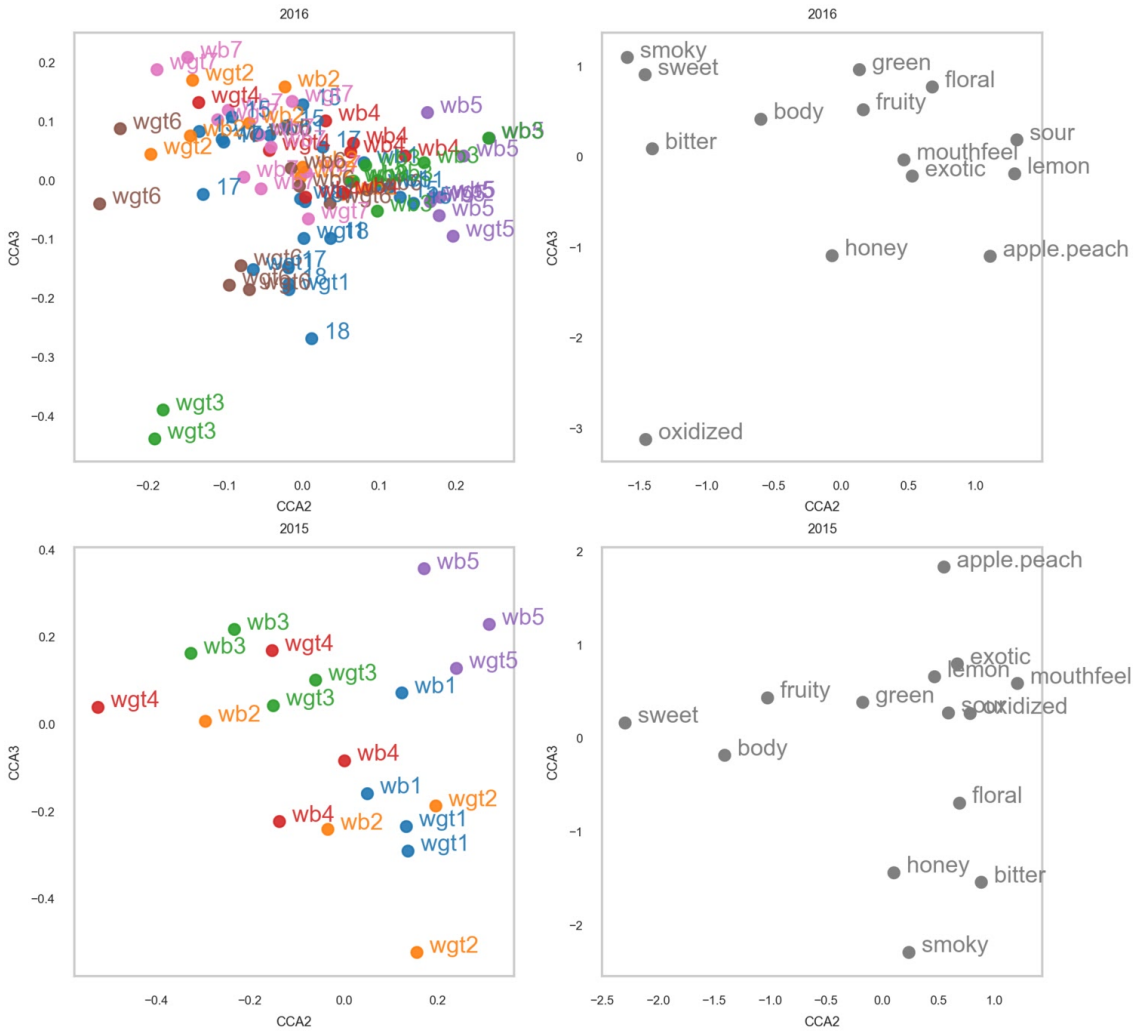


Figure S1. Constrained correspondence analysis (CCA) of the sensory analysis raw data with product and tasting replicate and variable scores visualised in the dimensions 2 and 3. Colors correspond to each vineyard: blue (vineyard 1), orange (vineyard 2), green (vineyard 3), red (vineyard 4), purple (vineyard 5), brown (vineyard 6) and pink (vineyard 7). wb = vineyard, wb = winery; numbers (1-7) corresponds to the each vineyards; 15 = vineyard 1 with SO_2; 17&18 = vineyard 1 with yeast 1 & yeast 2.

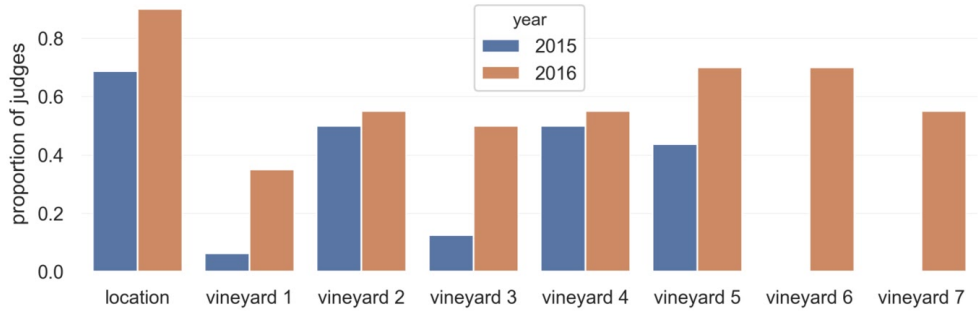


Figure S2. Proportion of judges who found important features for different targets in 2015 (blue) and 2016 (orange). Feature selected sensory attributes from each judge. The judges counts are scaled to corresponding panel sizes (2015: 16 judges, 2016: 20 judges). Location corresponds to the fermentation location.

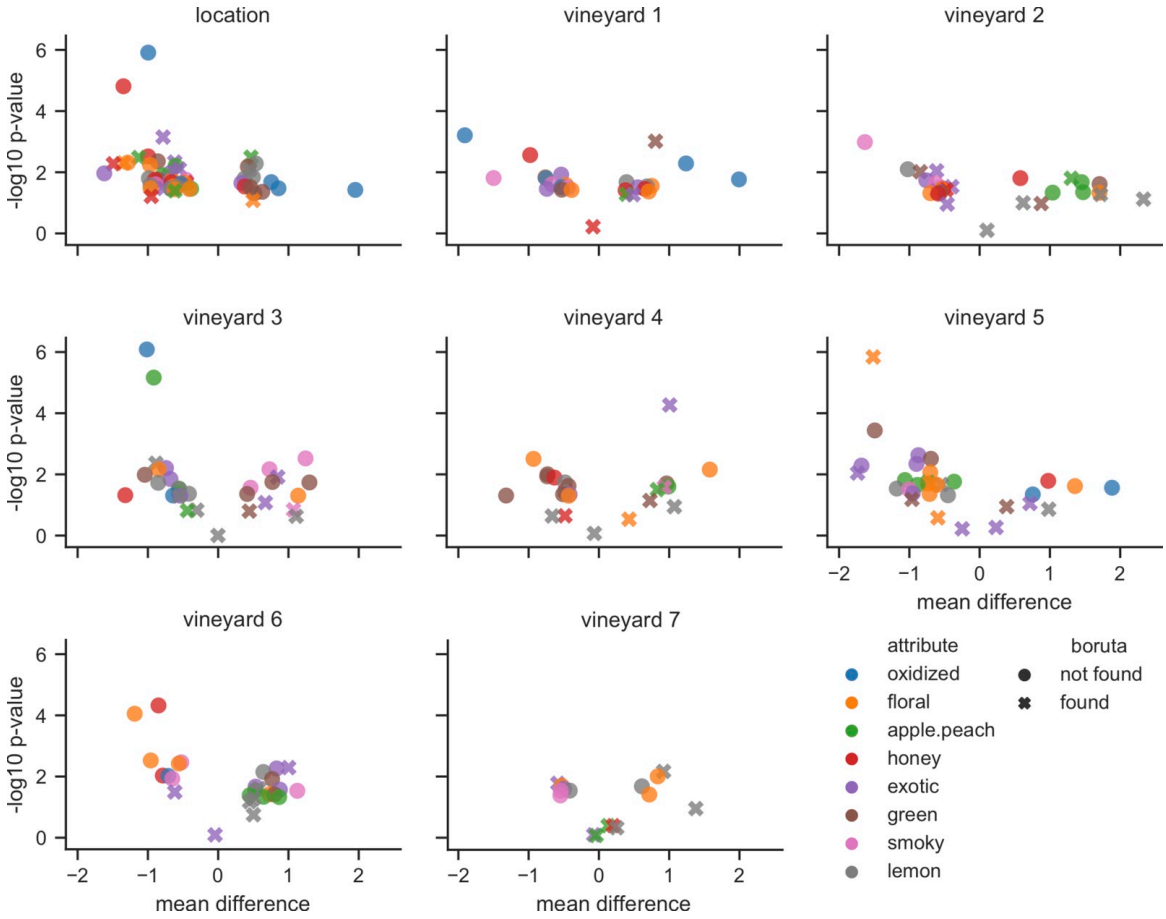
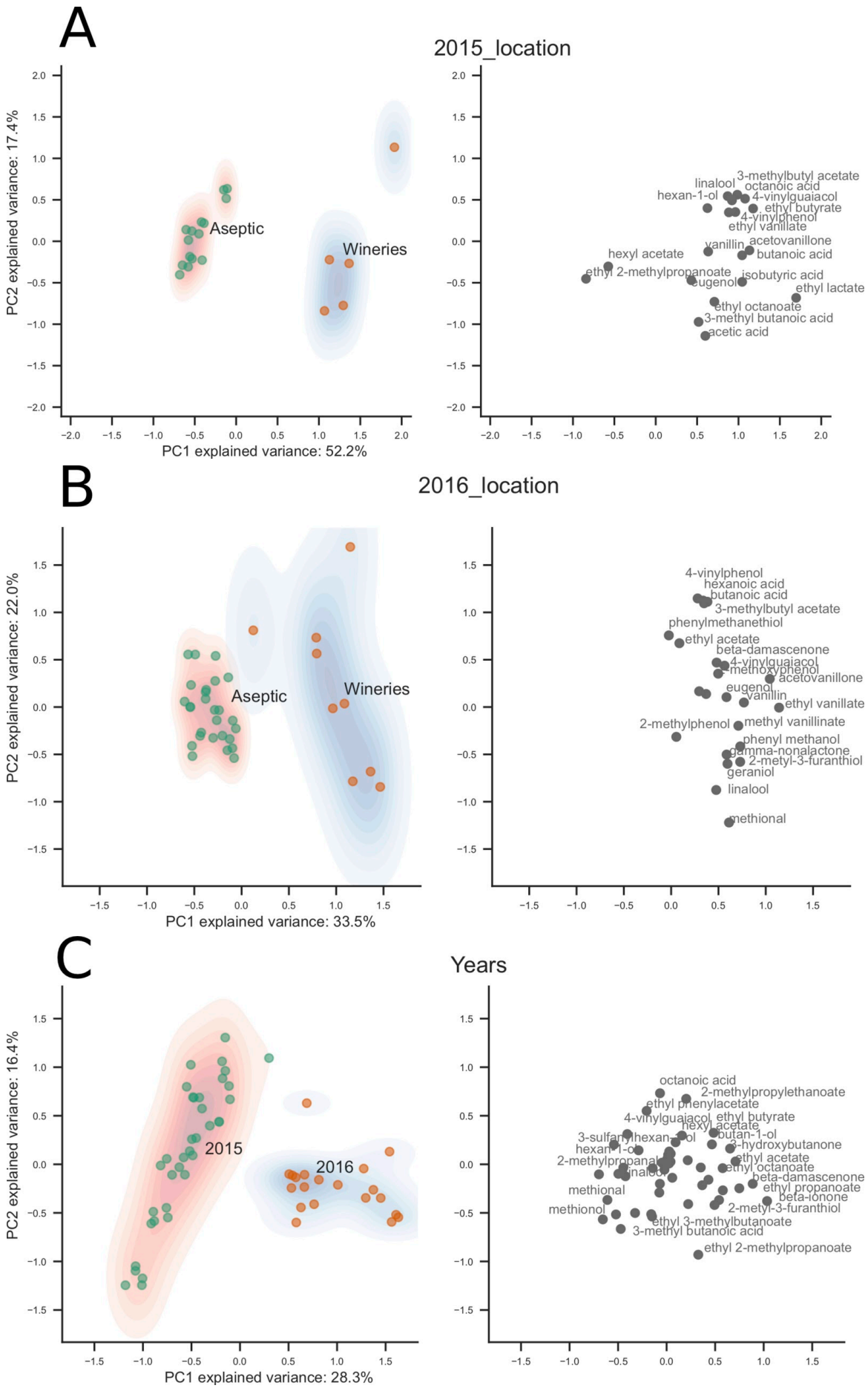


Figure S3. Olfactory attributes' mean difference changes and p-values for location and each vineyard in both years. The negative mean difference corresponds to: for location, mean values found higher in the winery than vineyards, and for vineyards, a higher mean values in the vineyard than others. The feature selection attributes are visualized with a cross, even if below discovery rate. Symbols: circle = found by feature selection and cross = not found by feature selection.


A
2015_location
PC2 explained variance: 17.4%
PC1 explained variance: 52.2%
Aseptic
Wineries
3-methylbutyl acetate
linalool
octanoic acid
hexan-1-ol
4-vinylguaiacol
ethyl butyrate
4-vinylphenol
ethyl vanillate
vanillin
acetovanillone
hexyl acetate
butanoic acid
ethyl 2-methylpropanoate
isobutyric acid
eugenol
ethyl lactate
ethyl octanoate
3-methyl butanoic acid
acetic acid
B
2016_location
PC2 explained variance: 22.0%
PC1 explained variance: 33.5%
Aseptic
Wineries
4-vinylphenol
hexanoic acid
butanoic acid
3-methylbutyl acetate
phenylmethanethiol
ethyl acetate
beta-damascenone
4-vinylguaiacol
4-methoxyphenol
acetovanillone
eugenol
vanillin
ethyl vanillate
2-methylphenol
methyl vanillinate
phenyl methanol
gamma-nonalactone
2-metyl-3-furanthiol
geraniol
linalool
methional
C
Years
PC2 explained variance: 16.4%
PC1 explained variance: 28.3%
2015
2016
octanoic acid
2-methylpropylethanoate
ethyl phenylacetate
4-vinylguaiacol
ethyl butyrate
hexyl acetate
3-sulfanylhexan-1-ol
butan-1-ol
hexan-1-ol
3-hydroxybutanone
2-methylpropanal
ethyl acetate
linalool
ethyl octanoate
beta-damascenone
methional
ethyl propanoate
beta-ionone
methionol
2-metyl-3-furanthiol
ethyl 3-methylbutanoate
3-methyl butanoic acid
ethyl 2-methylpropanoate

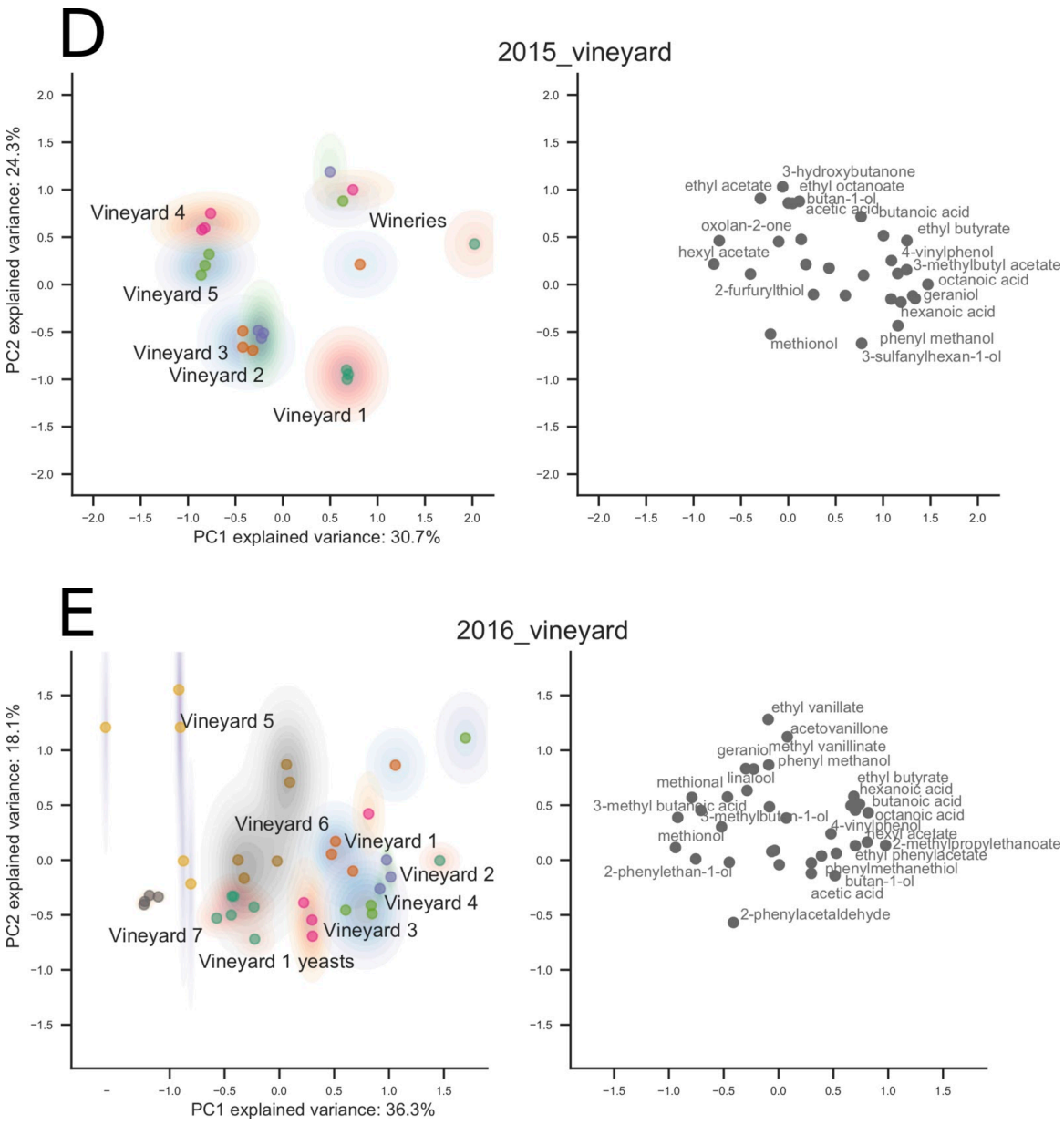
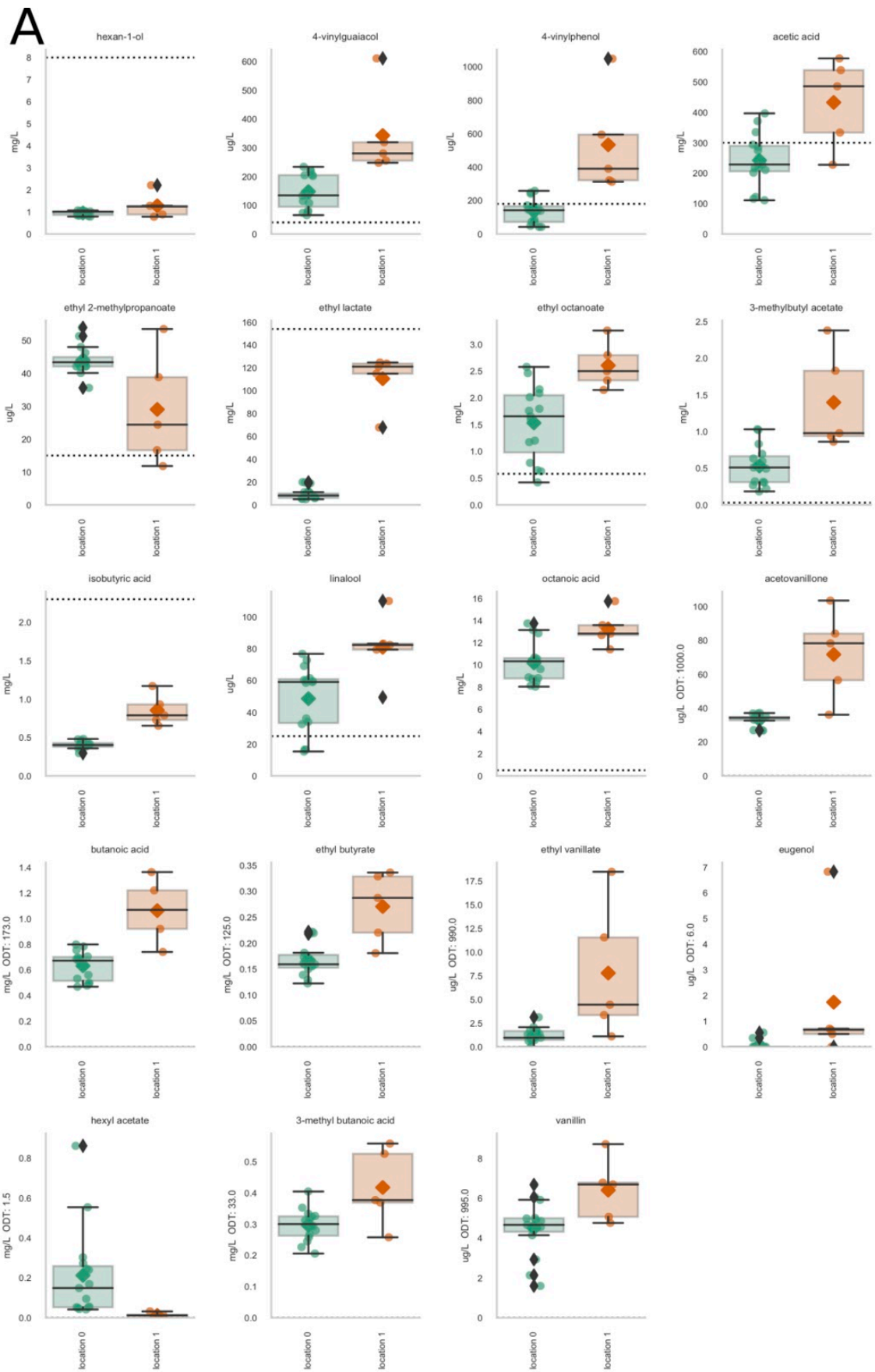
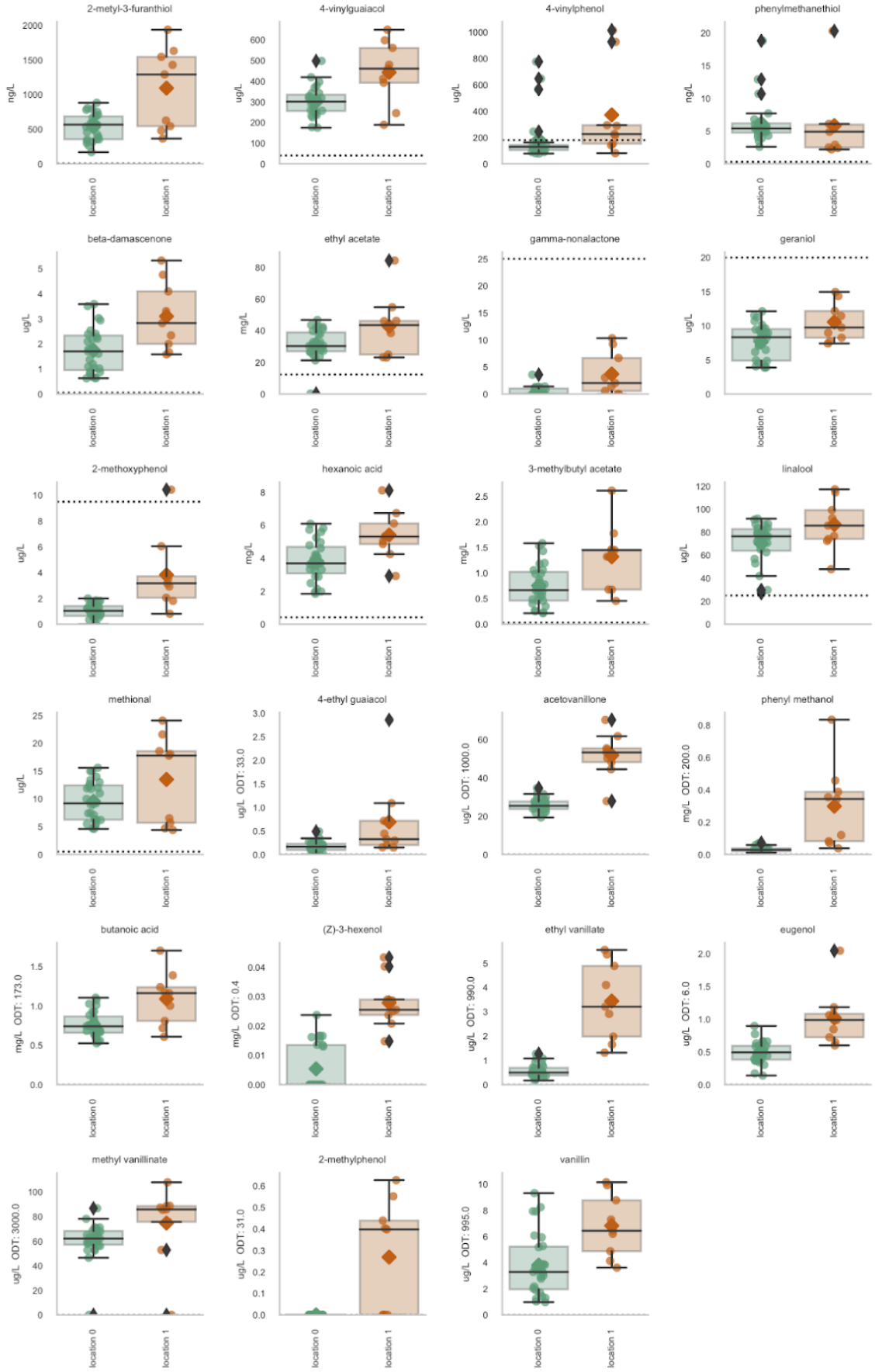


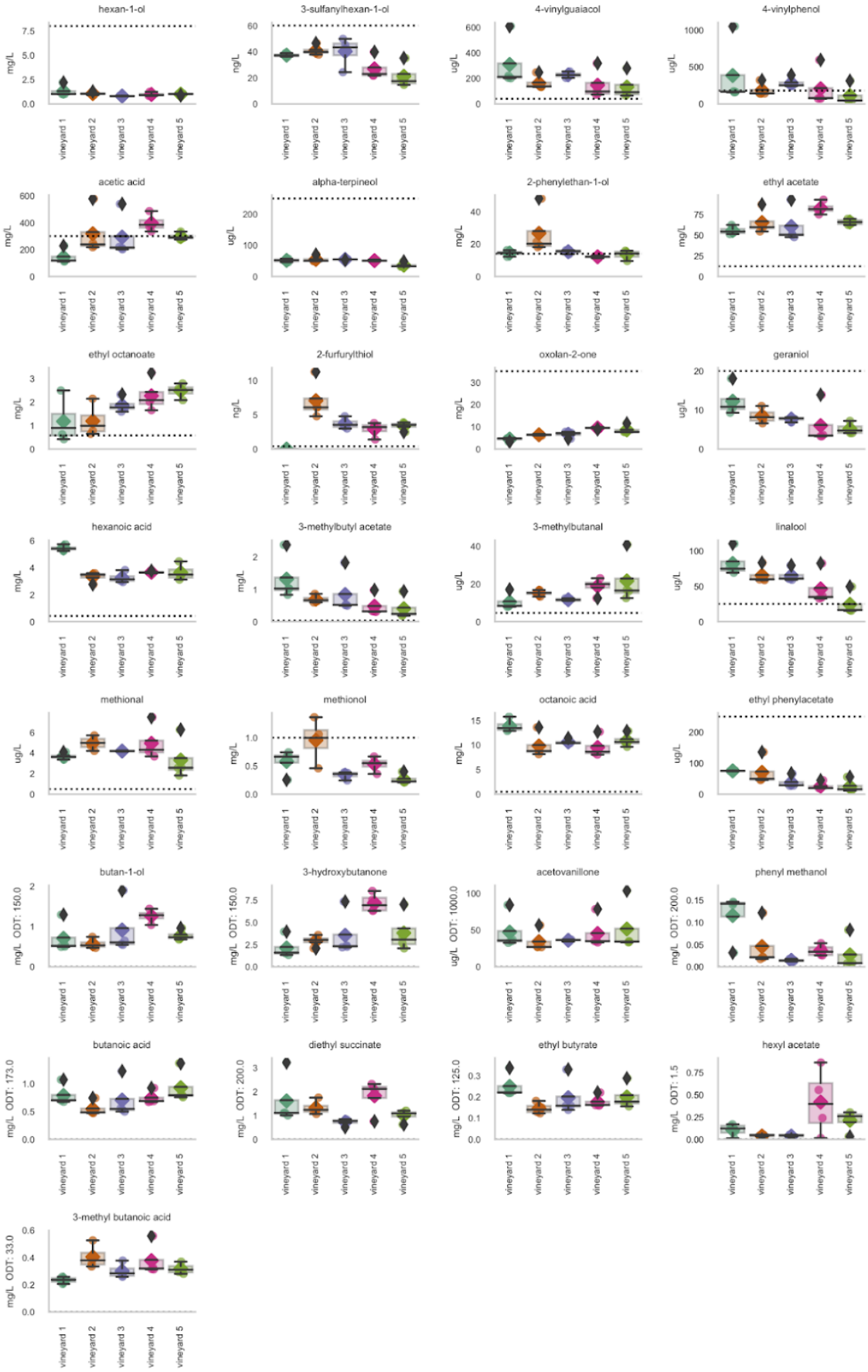
Figure S4. Random forest feature selected chemical compounds decomposed through principal component analysis. (A) 2015 fermentation location, (B) 2016 fermentation location, (C) years 2015 and 2016, (D) 2015 vineyards and (E) 2016 vineyards. Colours correspond to different classes. The shading corresponds to bivariate gaussian density estimation for the single class. Aseptic = vineyard in (A & B). Feature annotations that group close to the origo were removed for visualization purposes. See Figure S5 for all features.


A
hexan-1-ol
mg/L
4-vinylguaiacol
ug/L
4-vinylphenol
ug/L
acetic acid
mg/L
ethyl 2-methylpropanoate
ug/L
ethyl lactate
mg/L
ethyl octanoate
mg/L
3-methylbutyl acetate
mg/L
isobutyric acid
mg/L
linalool
ug/L
octanoic acid
mg/L
acetovanillone
ug/L ODT: 1000.0
butanoic acid
mg/L ODT: 173.0
ethyl butyrate
mg/L ODT: 125.0
ethyl vanillate
ug/L ODT: 990.0
eugenol
ug/L ODT: 6.0
hexyl acetate
mg/L ODT: 1.5
3-methyl butanoic acid
mg/L ODT: 33.0
vanillin
ug/L ODT: 995.0
location 0
location 1

B


2-metyl-3-furanthiol
ng/L
4-vinylguaiacol
ug/L
4-vinylphenol
ug/L
phenylmethanethiol
ng/L
beta-damascenone
ug/L
ethyl acetate
mg/L
gamma-nonalactone
ug/L
geraniol
ug/L
2-methoxyphenol
ug/L
hexanoic acid
mg/L
3-methylbutyl acetate
mg/L
linalool
ug/L
methional
ug/L
4-ethyl guaiacol
ug/L ODT: 33.0
acetovanillone
ug/L ODT: 1000.0
phenyl methanol
mg/L ODT: 200.0
butanoic acid
mg/L ODT: 173.0
(Z)-3-hexenol
mg/L ODT: 0.4
ethyl vanillate
ug/L ODT: 990.0
eugenol
ug/L ODT: 6.0
methyl vanillinate
ug/L ODT: 3000.0
2-methylphenol
ug/L ODT: 31.0
vanillin
ug/L ODT: 995.0
location 0
location 1

C


hexan-1-ol
mg/L
3-sulfanylhexan-1-ol
ng/L
4-vinylguaiacol
ug/L
4-vinylphenol
ug/L
acetic acid
mg/L
alpha-terpineol
ug/L
2-phenylethan-1-ol
mg/L
ethyl acetate
mg/L
ethyl octanoate
mg/L
2-furfurylthiol
ng/L
oxolan-2-one
mg/L
geraniol
ug/L
hexanoic acid
mg/L
3-methylbutyl acetate
mg/L
3-methylbutanal
ug/L
linalool
ug/L
methional
ug/L
methionol
mg/L
octanoic acid
mg/L
ethyl phenylacetate
ug/L
butan-1-ol
mg/L ODT: 150.0
3-hydroxybutanone
mg/L ODT: 150.0
acetovanillone
ug/L ODT: 1000.0
phenyl methanol
mg/L ODT: 200.0
butanoic acid
mg/L ODT: 173.0
diethyl succinate
mg/L ODT: 200.0
ethyl butyrate
mg/L ODT: 125.0
hexyl acetate
mg/L ODT: 1.5
3-methyl butanoic acid
mg/L ODT: 33.0
vineyard 1
vineyard 2
vineyard 3
vineyard 4
vineyard 5

D

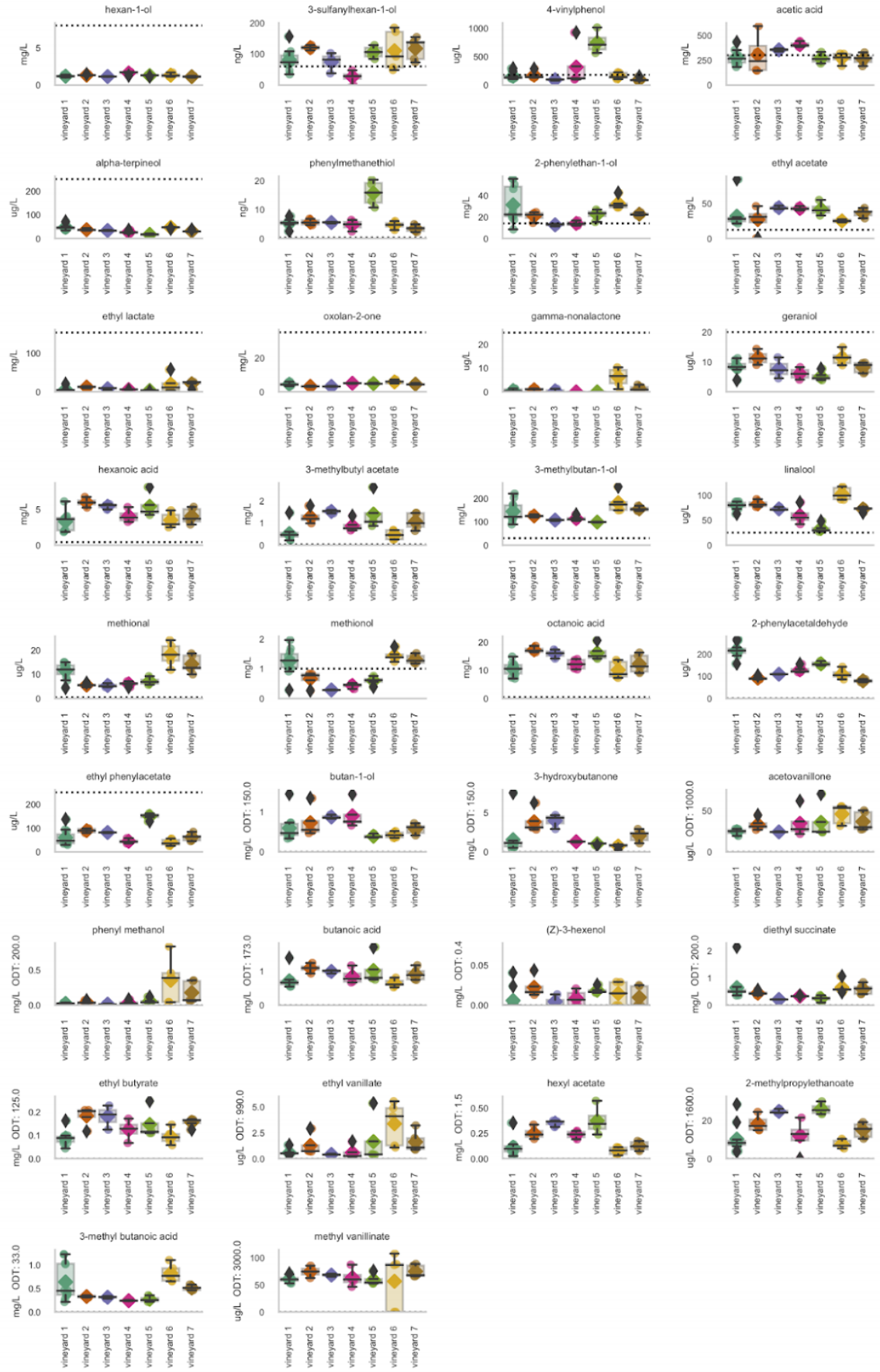

hexan-1-ol
mg/L
3-sulfanylhexan-1-ol
ng/L
4-vinylphenol
ug/L
acetic acid
mg/L
alpha-terpineol
ug/L
phenylmethanethiol
ng/L
2-phenylethan-1-ol
mg/L
ethyl acetate
mg/L
ethyl lactate
mg/L
oxolan-2-one
mg/L
gamma-nonalactone
ug/L
geraniol
ug/L
hexanoic acid
mg/L
3-methylbutyl acetate
mg/L
3-methylbutan-1-ol
mg/L
linalool
ug/L
methional
ug/L
methionol
mg/L
octanoic acid
mg/L
2-phenylacetaldehyde
ug/L
ethyl phenylacetate
ug/L
butan-1-ol
mg/L ODT: 150.0
3-hydroxybutanone
mg/L ODT: 150.0
acetovanillone
ug/L ODT: 1000.0
phenyl methanol
mg/L ODT: 200.0
butanoic acid
mg/L ODT: 173.0
(Z)-3-hexenol
mg/L ODT: 0.4
diethyl succinate
mg/L ODT: 200.0
ethyl butyrate
mg/L ODT: 125.0
ethyl vanillate
ug/L ODT: 990.0
hexyl acetate
mg/L ODT: 1.5
2-methylpropylethanoate
ug/L ODT: 1600.0
3-methyl butanoic acid
mg/L ODT: 33.0
methyl vanillinate
ug/L ODT: 3000.0
vineyard 1
vineyard 2
vineyard 3
vineyard 4
vineyard 5
vineyard 6
vineyard 7

Figure S5. Boxplots of feature selected sensory attributes with their corresponding

Chemical name	Odour threshold	Detection limit	Reported in range	InChiKey	CAS-nr	FooDB id	Superclass	Class	Subclass	Parent Level 1
acetaldehyde	0.5	115.00	mg/L	IKHGUXGNUITLKF-UHFFFAOYSA-N	75--07-0	FDB008297	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Aldehydes
acetic acid	300	0.24	mg/L	QTBSBXVTEAMEQO-UHFFFAOYSA-N	64--19--7	FDB019725	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acids	Carboxylic acids
3-hydroxybutanone	150	0.10	mg/L	ROWKJAVDOGWPAT-UHFFFAOYSA-N	513--86--0	FDB011799	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Acyloins
phenyl methanol	200	0.01	mg/L	WVDDGKGOMKODPV-UHFFFAOYSA-N	100-51-6	FDB008745	Benzenoids	Benzene and substituted derivatives	Benzyl alcohols	Benzyl alcohols
2-phenylethan-1-ol	14	5.00	mg/L	WRMNZCZEMHIOCP-UHFFFAOYSA-N	60--12--8	FDB012152	Benzenoids	Benzene and substituted derivatives	Benzyl alcohols	Benzene and substituted derivatives
butanoic acid	173	0.10	mg/L	FERIUCNNQQJTOY-UHFFFAOYSA-N	107--92--6	FDB012062	Lipids and lipid-like molecules	Fatty Acyls	Fatty acids and conjugates	Straight chain fatty acids
(Z)-3-hexenol	0.4	0.01	mg/L	UFLHIIWVXFIJGU-ARJAWSKDSA-N	928--96--1	FDB008091	Lipids and lipid-like molecules	Fatty Acyls	Fatty alcohols	Fatty alcohols
decanoic acid	1	27.00	mg/L	GHVNFZFCNZKVNT-UHFFFAOYSA-N	334--48--5	FDB012027	Lipids and lipid-like molecules	Fatty Acyls	Fatty acids and conjugates	Medium-chain fatty acids
diacetyl	0.1	25.00	mg/L	QSJXEFYPDANLFS-UHFFFAOYSA-N	431--03--8	FDB011930	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Ketones
diethyl succinate	200	3.00	mg/L	DKMROQRQHGEIOW-UHFFFAOYSA-N	123-25-1	FDB012012	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl acetate	12.3	0.01	mg/L	XEKOWRVHYACXOJ-UHFFFAOYSA-N	141--78--6	FDB003240	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
ethyl butyrate	125	0.03	mg/L	OBNCKNCVKJNDBV-UHFFFAOYSA-N	105--54--4	FDB012082	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl decanoate	0.2	0.02	mg/L	RGXWDWUGBUHDO-UHFFFAOYSA-N	110--38--3	FDB002986	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl hexanoate	62	0.03	mg/L	SHZIWNPUGXLXDT-UHFFFAOYSA-N	123--66--0	FDB019921	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl lactate	154	0.10	mg/L	LZCLXQDLBQLTDK-UHFFFAOYSA-N	97-64-3	FDB020545	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
ethyl octanoate	0.58	12.00	mg/L	YYZUSRORWSJGET-UHFFFAOYSA-N	106--32--1	FDB019907	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl propanoate	5.5	0.05	mg/L	FKRCODPIKNYEAC-UHFFFAOYSA-N	105-37-3	FDB001366	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
oxolan-2-one	35	0.02	mg/L	YEJRWHAVMIAJKC-UHFFFAOYSA-N	96-48-0	FDB003392	Organoheterocyclic compounds	Lactones	Gamma butyrolactones	Gamma butyrolactones
hexanoic acid	0.42	0.01	mg/L	FUZZWVXGSFPDMH-UHFFFAOYSA-N	142--62--1	FDB013897	Lipids and lipid-like molecules	Fatty Acyls	Fatty acids and conjugates	Medium-chain fatty acids
hexyl acetate	1.5	0.01	mg/L	AOGQPLXWSUTHQB-UHFFFAOYSA-N	142--92--7	FDB001267	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
3-methylbutyl acetate	0.03	18.00	mg/L	MLFHJEHSLIIPHL-UHFFFAOYSA-N	123--92--2	FDB008132	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
3-methylbutan-1-ol	30	0.02	mg/L	PHTQWCKDNZKARW-UHFFFAOYSA-N	123--51--3	FDB008131	Organic oxygen compounds	Organooxygen compounds	Alcohols and polyols	Primary alcohols
2-methylpropan-1-ol	40	0.02	mg/L	ZXEKIBDNHEJCQA-UHFFFAOYSA-N	78--83--1	FDB003274	Organic oxygen compounds	Organooxygen compounds	Alcohols and polyols	Primary alcohols
isobutyric acid	2.3	0.10	mg/L	KQNPFQTWMSNSAP-UHFFFAOYSA-N	79--31--2	FDB003277	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acids	Carboxylic acids
3-methyl butanoic acid	33	28.00	mg/L	GWYFCOCPABKNJV-UHFFFAOYSA-N	503-74-2	FDB001324	Lipids and lipid-like molecules	Fatty Acyls	Fatty acids and conjugates	Branched fatty acids
methionol	1	26.00	mg/L	CZUGFKJYCPYHHV-UHFFFAOYSA-N	505--10--2	FDB008379	Organosulfur compounds	Thioethers	Dialkylthioethers	Dialkylthioethers
octanoic acid	0.5	0.01	mg/L	WWZKQHOCKIZLMA-UHFFFAOYSA-N	124--07--2	FDB003336	Lipids and lipid-like molecules	Fatty Acyls	Fatty acids and conjugates	Medium-chain fatty acids
butan-1-ol	150	2.00	mg/L	LRHPLDYGYMQRHN-UHFFFAOYSA-N	71--36--3	FDB012814	Organic oxygen compounds	Organooxygen compounds	Alcohols and polyols	Primary alcohols
hexan-1-ol	8	14.00	mg/L	ZSIAUFGUXNUGDI-UHFFFAOYSA-N	111--27--3	FDB008072	Lipids and lipid-like molecules	Fatty Acyls	Fatty alcohols	Fatty alcohols
2-methylbutanal	16	4.00	µg/L	BYGQBDHUGHBGMD-UHFFFAOYSA-N	96--17--3	FDB008127	Organic oxygen compounds	Organooxygen compounds	Aldehydes	Aldehydes
2,6-dimethoxyphenol	570	0.07	µg/L	KLIDCXVFHGNTTM-UHFFFAOYSA-N	91--10--1	FDB012443	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
4-ethyl guaiacol	33	0.02	µg/L	CHWNEIVBYREQRF-UHFFFAOYSA-N	2785--89--9	FDB019887	Benzenoids	Phenols	Monoterpenoids	Methoxyphenols
4-ethyl phenol	35	0.02	µg/L	HXDOZKJGKXYMEW-UHFFFAOYSA-N	123--07--9	FDB000358	Benzenoids	Phenols	1-hydroxy-2-unsubstituted benzenoids	1-hydroxy-2-unsubstituted benzenoids
4-vinylguaiacol	40	0.02	µg/L	YOMSJEATGXXYPX-UHFFFAOYSA-N	7786--61--0	FDB000857	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
4-vinylphenol	180	0.06	µg/L	?UGYGGDSWSUORM-UHFFFAOYSA-N	2628--17--3	FDB010540	Benzenoids	Benzene and substituted derivatives	Styrenes	Styrenes
acetovanillone	1000	0.14	µg/L	ZFFTZDQKIXPDAF-UHFFFAOYSA-N	498--02--2	FDB010898	Organoheterocyclic compounds	Heteroaromatic compounds	Aldehydes	Heteroaromatic compounds
alpha-ionone	2.6	0.33	µg/L	UZFLPKAIBPNNCA-BQYQJAHWSA-N	127-41-3	FDB014484	Lipids and lipid-like molecules	Prenol lipids	Sesquiterpenoids	Sesquiterpenoids
alpha-terpineol	250	0.05	µg/L	WUOACPNHFRMFPN-UHFFFAOYSA-N	98--55--5	FDB014922	Lipids and lipid-like molecules	Prenol lipids	Monoterpenoids	Menthane monoterpenoids
benzaldehyde	2000	0.03	µg/L	HUMNYLRZRPPJDN-UHFFFAOYSA-N	100-52-7	FDB014661	Benzenoids	Benzene and substituted derivatives	Benzoyl derivatives	Benzoyl derivatives
beta-damascenone	0.05	0.19	µg/L	POJARNZEYGURDG-FNORWQNLSA-N	23726--93--4	FDB015479	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Alpha,beta-unsaturated carbonyl compounds
beta-ionone	0.09	0.33	µg/L	PSQYTAPXSHCGMF-BQYQJAHWSA-N	79--77--6	FDB015469	Lipids and lipid-like molecules	Prenol lipids	Sesquiterpenoids	Sesquiterpenoids
butyl acetate	1800	0.17	µg/L	DKPFZGUDAPQIHT-UHFFFAOYSA-N	123--86--4	FDB003386	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
c-whiskylactone	67	0.33	µg/L	WNVCMFHPRIBNCW-UHFFFAOYSA-N	55013--32--6	FDB029743	Organoheterocyclic compounds	Lactones	Gamma butyrolactones	Gamma butyrolactones
citronellol	100	0.78	µg/L	QMVPMAAFGQKVCJ-UHFFFAOYSA-N	106--22--9	FDB014490	Lipids and lipid-like molecules	Prenol lipids	Monoterpenoids	Acyclic monoterpenoids
trans-isoeugenol	6	0.04	µg/L	BJIOGJUNALELMI-ONEGZZNKSA-N	97--54--1	FDB012403	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
ethyl 2-methylbutanoate	18	0.33	µg/L	HCRBXQFHJMCTLF-UHFFFAOYSA-N	7452--79--1	FDB029639	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl cinnamate	1.1	0.33	µg/L	KBEBGUQPQBELIU-HJWRWDBZSA-N	103--36--6	FDB012002	Phenylpropanoids and polyketides	Cinnamic acids and derivatives	Cinnamic acid esters	Cinnamic acid esters
ethyl dihydrocinnamate	1.6	0.03	µg/L	QDCUBAFTOOPBFR-UHFFFAOYSA-N	2021-28-5	FDB016196	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl 2-methylpropanoate	15	495.00	µg/L	WDAXFOBOLVPGLV-UHFFFAOYSA-N	97--62--1	FDB003278	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
ethyl 3-methylbutanoate	3	0.33	µg/L	PPXUHEORWJQRHJ-UHFFFAOYSA-N	3301-94-8	FDB001326	Lipids and lipid-like molecules	Fatty Acyls	Fatty acid esters	Fatty acid esters
ethyl vanillate	990	0.06	µg/L	ULKIVNGMBPJKRW-UHFFFAOYSA-N	617-05-0	FDB029772	Benzenoids	Benzene and substituted derivatives	Benzoic acids and derivatives	Methoxybenzoic acids and derivatives
eugenol	6	0.33	µg/L	RRAFCDWBNXTKKO-UHFFFAOYSA-N	97--53--0	FDB012171	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
gamma-decalactone	0.7	0.33	µg/L	IPYYFLINQYPWGJ-UHFFFAOYSA-N	706--14--9	FDB008393	Organoheterocyclic compounds	Lactones	Gamma butyrolactones	Gamma butyrolactones
gamma-nonalactone	25	0.06	µg/L	OALYTRUKMRCXNH-UHFFFAOYSA-N	104--61--0	FDB008395	Organoheterocyclic compounds	Lactones	Gamma butyrolactones	Gamma butyrolactones
geraniol	20	0.33	µg/L	GLZPCOQZEFWAFX-JXMROGBWSA-N	106--24--1	FDB013792	Lipids and lipid-like molecules	Prenol lipids	Monoterpenoids	Acyclic monoterpenoids
2-methoxyphenol	9.5	0.05	µg/L	LHGVFZTZFXWLCP-UHFFFAOYSA-N	90--05--1	FDB011885	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
2-methylpropyl ethanoate	1600	0.16	µg/L	GJRQTCIYDGXPES-UHFFFAOYSA-N		FDB003275	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
2-methylpropanal	6	0.33	µg/L	AMIMRNSIRUDHCM-UHFFFAOYSA-N	78--84--2	FDB003271	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Aldehydes
3-methylbutanal	4.6	0.98	µg/L	YGHRJJRRZDOVPD-UHFFFAOYSA-N	590-86-3	FDB003285	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Aldehydes
linalool	25	0.04	µg/L	CDOSHBSSFJOMGT-UHFFFAOYSA-N	78--70--6	FDB014940	Lipids and lipid-like molecules	Prenol lipids	Monoterpenoids	Acyclic monoterpenoids
3-methylphenol	68	0.00	µg/L	RLSSMJSEOOYNOY-UHFFFAOYSA-N	108--39--4	FDB008788	Benzenoids	Phenols	Cresols	Meta cresols
methional	0.5	0.12	µg/L	CLUWOWRTHNBBIJ-UHFFFAOYSA-N	3268--49--3	FDB008541	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Aldehydes
methyl vanillinate	3000	0.04	µg/L	BVWTXUYLKBHMOX-UHFFFAOYSA-N	3943-74-6	FDB029771	Benzenoids	Benzene and substituted derivatives	Benzoic acids and derivatives	Methoxybenzoic acids and derivatives
2-methylphenol	31	0.33	µg/L	JWVGKYWNOKOFNN-UHFFFAOYSA-N	95--48--7	FDB008785	Benzenoids	Phenols	Cresols	Ortho cresols
2-phenylacetaldehyde	1	3.00	µg/L	DTUQWGWMVIHBKE-UHFFFAOYSA-N	122--78--1	FDB012238	Benzenoids	Benzene and substituted derivatives	Phenylacetaldehydes	Phenylacetaldehydes
ethyl phenylacetate	250	0.02	µg/L	DULCUDSUACXJJC-UHFFFAOYSA-N	101--97--3	FDB010560	Benzenoids	Benzene and substituted derivatives	Benzoic acid derivatives	Benzene and substituted derivatives
trans-whiskylactone	790	0.09	µg/L	WNVCMFHPRIBNCW-UHFFFAOYSA-N	105119--22--0	FDB014731	Organoheterocyclic compounds	Lactones	Gamma butyrolactones	Gamma butyrolactones
vanillin	995	0.08	µg/L	MWOOGOJBHIARFG-UHFFFAOYSA-N	121--33--5	FDB000838	Benzenoids	Phenols	Methoxyphenols	Methoxyphenols
2-metyl-3-furanthiol	4	0.30	ng/L	SRUTWBWLFKSTIS-UHFFFAOYSA-N	65505--17--1	FDB019911	Organoheterocyclic compounds	Heteroaromatic compounds		Heteroaromatic compounds
3-sulfanylhexan-1-ol	60	6.60	ng/L	TYZFMFVWHZKYSE-UHFFFAOYSA-N	51755-83-0	FDB019850	Organosulfur compounds	Thiols	Alkylthiols	Alkylthiols
3-sulfanylhexyl acetate	4	1.20	ng/L	JUCARGIKESIVLB-UHFFFAOYSA-N	136954--20--6	FDB018425	Organic acids and derivatives	Carboxylic acids and derivatives	Carboxylic acid derivatives	Carboxylic acid esters
4-methyl-4-sulfanylpentan-2-one	0.8	0.80	ng/L	QRNZMFDCKKEPSX-UHFFFAOYSA-N	19872--52--7	FDB008119	Organic oxygen compounds	Organooxygen compounds	Carbonyl compounds	Ketones
phenylmethanethiol	0.3	0.20	ng/L	UENWRTRMJIOCKN-UHFFFAOYSA-N	100-53-8	FDB000803	Benzenoids	Benzene and substituted derivatives		Benzene and substituted derivatives
2-furfurylthiol	0.4	0.20	ng/L	ZFFTZDQKIXPDAF-UHFFFAOYSA-N	98--02--2	FDB010898	Organoheterocyclic compounds	Heteroaromatic compounds		Heteroaromatic compounds

sensory threshold. (A) 2015 fermentation location, (B) 2016 fermentation location, (C)

2015 vineyards and (D) 2016 vineyards. Location 0 represented vineyard and location 1 represented winery. For visualization purposes, if the values were significantly under sensory threshold, the threshold has been reported in the y-axis label instead of the line drawn.

Sheet13

	butan-1-ol	hexan-1-ol	2-methylbutanal	2-methyl-3-furanthiol	2,6-dimethoxyphenol	3-sulfanylhexan-1-ol	3-sulfanylhexyl acetate	4-ethyl guaiacol	4-ethyl phenol	4-methyl-4-sulfanylpentan-2-one	4-vinylguaiacol	4-vinylphenol	acetaldehyde	acetic acid
wb1a_2016	1.7920E+01	2.8487E+00	3.6923E-03	3.5220E-04	1.7008E-01	1.5630E-04	2.2500E-06	1.5978E-02	2.3236E-02	5.6000E-06	2.8913E-01	4.7160E-02	5.5294E-02	4.4028E+01
wb1b_2016	4.6888E+01	4.0266E+00	6.5177E-03	6.7920E-04	5.1922E-01	7.0300E-05	3.5790E-04	2.1265E-02	2.8454E-03	8.3600E-08	3.9679E-01	6.1944E-02	4.0410E-01	2.6295E+02
wb1c_2016	1.4264E+02	2.3343E+00	1.4375E-02	5.8330E-04	4.5541E-01	7.7100E-05	2.8610E-04	1.5010E-02	2.5038E-02	3.8500E-07	3.4174E-01	1.1068E-02	9.4806E-02	3.0129E+02
wb2a_2016	5.9505E+00	2.5013E-01	3.1401E-03	3.2790E-04	2.4063E-01	1.2320E-04	4.1000E-06	6.5675E-03	2.4108E-02	2.9000E-06	3.0930E-01	1.7691E-01	4.4138E-01	2.4944E+02
wb2b_2016	4.9448E+01	2.5362E+00	3.8840E-03	5.4370E-04	3.4280E-01	1.1150E-04	7.5100E-05	2.3389E-02	3.1501E-02	3.5000E-06	4.9784E-01	1.1355E-01	5.2595E-01	5.9188E+02
wb2c_2016	1.0957E+02	4.5174E+00	5.6941E-03	3.1690E-04	3.6054E-01	1.1800E-04	2.9300E-06	6.9925E-03	3.0881E-02	4.4000E-06	3.0071E-01	2.1219E-02	2.0047E-01	2.7799E+02
wb3a_2016	1.0044E+02	2.0365E+00	1.0028E-02	5.8370E-04	2.8540E-01	8.2000E-05	3.4200E-06	3.0497E-02	3.4173E-02	3.1400E-07	1.7436E-01	1.5265E-01	1.0957E-02	3.6766E+02
wb3b_2016	3.3419E+01	6.6289E+00	1.3123E-03	3.4860E-04	7.7777E-02	1.0250E-04	3.0400E-06	2.6109E-02	1.4449E-02	3.9200E-07	1.7729E-01	9.6831E-02	4.0930E-01	3.3682E+02
wb3c_2016	1.1061E+02	6.8217E+00	2.5912E-03	3.5610E-04	1.8114E-01	1.1500E-05	1.2800E-07	2.1402E-02	2.4484E-02	6.0400E-07	2.4418E-01	4.0527E-02	1.1595E-01	3.5551E+02
wb4a_2016	2.6704E+01	3.5599E+00	7.1300E-03	6.9320E-04	1.4685E-02	5.8700E-05	1.1570E-04	1.0260E-02	2.0877E-02	1.8500E-07	2.3728E-01	1.5327E-01	6.5818E-02	4.4188E+02
wb4b_2016	1.0760E+02	6.7238E+00	1.1549E-02	7.0750E-04	4.0536E-01	2.7100E-05	3.1870E-04	7.0838E-03	1.9971E-02	4.9660E-07	2.5970E-01	1.6749E-01	2.5376E+00	3.8258E+02
wb4c_2016	1.3160E+02	3.7916E+00	7.9728E-03	7.4960E-04	2.1450E-01	3.7500E-05	6.0800E-07	6.7343E-03	1.8951E-02	4.4200E-07	2.5408E-01	9.7435E-02	4.3739E-01	3.8740E+02
wb5a_2016	8.0937E+01	4.4428E+00	1.7239E-03	5.6780E-04	1.1963E-01	9.4600E-05	2.6500E-05	7.4480E-03	7.0848E-04	1.9300E-07	3.6775E-01	6.4668E-01	2.8675E-01	1.7338E+02
wb5b_2016	6.2259E+01	7.7132E+00	4.9474E-03	5.3590E-04	1.9613E-01	1.1770E-04	3.3700E-06	3.2187E-02	2.4520E-02	4.9900E-07	3.1958E-01	5.6547E-01	1.9209E-03	2.1844E+02
wb5c_2016	1.4706E+02	5.9864E+00	1.2990E-03	2.9890E-04	1.3717E-01	1.2830E-04	3.3600E-06	2.3131E-02	2.9678E-02	1.1400E-07	3.0460E-01	7.7556E-01	2.9077E-01	2.2359E+02
wb6a_2016	1.2553E+02	4.0643E+00	9.5117E-04	4.2290E-04	1.5118E-01	1.8450E-04	2.2800E-06	1.2803E-02	7.9947E-03	3.0700E-07	2.7731E-01	1.1733E-01	2.2579E-01	3.1566E+02
wb6c_2016	4.6565E+01	1.3755E+00	9.8476E-03	7.9000E-04	3.7051E-01	1.7110E-04	2.6370E-04	1.9664E-02	2.8043E-02	4.5300E-07	3.2771E-01	7.1680E-03	3.2481E-01	3.1309E+02
wb7a_2016	8.9469E+01	7.8752E+00	4.1149E-03	5.6560E-04	3.8246E-01	1.5430E-04	2.4760E-04	2.3050E-02	7.2522E-03	4.7800E-07	4.1894E-01	1.3740E-01	7.0362E-02	2.8094E+02
wb7b_2016	1.0864E+02	5.0343E+00	7.9900E-03	6.0520E-04	4.6551E-01	1.3970E-04	2.7100E-04	1.2558E-02	1.5557E-02	1.3300E-07	3.7794E-01	7.3203E-02	1.7912E-01	3.2774E+02
wb7c_2016	1.2984E+02	1.5260E+00	1.2528E-02	6.6510E-04	4.6223E-01	1.3700E-04	1.2300E-06	6.5539E-03	1.3537E-02	5.4800E-07	2.3238E-01	1.5888E-01	4.4826E-01	1.7616E+02
wgt1_2016	1.8613E+01	4.8488E-01	1.2320E-02	4.7930E-04	5.6306E-01	9.6900E-05	1.1740E-04	1.4901E-03	2.5036E-02	4.3800E-07	3.9224E-01	2.9350E-01	1.0425E-01	4.3178E+02
wgt2_2016	3.9754E+01	3.5554E+00	3.6124E-03	5.4320E-04	1.4807E-01	1.2860E-04	2.7500E-06	2.4880E-02	2.8178E-02	3.2600E-07	6.4880E-01	8.8335E-01	2.9407E-01	3.2832E+02
wgt4_2016	8.7063E+01	7.7898E+00	7.3735E-03	6.2450E-04	5.5796E-01	4.3600E-05	7.8000E-07	2.8064E-02	1.8542E-02	8.4800E-08	5.9789E-01	9.2598E-01	4.4779E-01	4.0950E+02
wgt5_2016	7.7623E+01	6.4635E+00	7.0700E-05	3.6380E-04	4.9909E-01	8.3800E-05	3.1700E-05	7.5585E-03	2.8535E-02	1.2900E-07	4.7930E-01	1.0135E+00	4.2814E-01	3.3164E+02
wgt6f77_2016	8.0753E+01	7.9851E+00	1.0578E-02	1.6302E-03	2.1507E-01	2.0800E-05	1.3299E-03	5.0085E-03	8.6956E-03	3.8600E-07	2.4518E-01	7.5732E-02	3.8214E-02	1.8406E+02
wgt6f775_2016	2.2320E+01	3.1550E+00	1.4070E-02	1.4280E-03	2.9439E-01	9.2000E-05	1.2400E-06	2.6215E-02	1.8631E-02	3.0600E-07	4.1154E-01	2.1683E-01	3.0531E-01	2.8062E+02
wgt6f202_2016	7.6403E+01	6.1048E+00	8.7286E-03	1.5444E-03	3.2469E-01	1.9900E-05	9.1300E-07	2.1403E-02	1.1716E-03	3.5900E-07	5.6096E-01	2.2557E-01	3.0983E-01	2.1200E+01
wgt7f209_2016	9.3716E+01	1.7536E+00	3.5259E-03	1.9363E-03	3.9917E-01	7.2900E-05	1.1720E-04	1.4796E-02	2.0867E-02	2.7400E-07	4.6059E-01	1.1594E-01	3.9357E-01	1.5353E+02
wgt7f30_2016	3.4699E+01	5.0526E+00	1.1976E-02	1.2899E-03	3.9247E-02	8.3700E-05	1.2690E-04	2.7794E-02	6.2571E-03	2.7600E-07	1.8844E-01	9.6399E-02	4.4599E-01	6.6043E+01
15a_2016	1.1799E+02	3.9370E+00	1.2752E-02	2.6200E-04	2.3326E-01	9.4100E-05	1.9440E-04	9.7530E-03	3.1891E-02	1.2400E-07	3.4173E-01	2.4484E-01	4.6187E-01	5.5577E+01
15b_2016	1.4139E+02	4.0147E+00	5.1402E-03	1.6810E-04	1.3804E-01	6.6200E-05	1.0900E-06	1.8171E-02	3.0893E-02	5.3300E-07	3.1400E-01	7.4820E-02	8.5400E-02	9.0889E+01
15c_2016	4.5589E+01	2.1157E+00	7.9628E-04	8.0370E-04	1.0742E-01	1.0850E-04	3.4600E-05	2.3572E-02	2.0258E-02	4.6200E-07	2.5910E-01	1.2009E-01	4.3517E-01	3.5096E+02
17a_2016	4.4786E+01	3.5201E+00	1.1576E-03	4.5700E-04	1.3639E-01	7.2600E-05	1.3600E-07	3.1236E-02	3.4000E-02	5.1500E-07	2.5701E-01	2.8116E-02	2.3856E-01	1.9427E+02
17b_2016	1.4359E+02	4.2483E+00	1.3705E-02	8.8120E-04	1.6261E-01	7.3400E-05	1.0290E-04	3.1000E-02	1.9470E-03	4.8000E-07	2.9899E-01	7.4688E-02	3.6840E-01	1.0610E+01
18a_2016	5.8217E+01	6.1448E-01	1.2102E-02	3.8700E-04	3.8347E-01	1.8500E-06	2.6800E-05	6.5264E-03	2.5071E-02	5.4800E-08	2.2468E-01	1.7080E-01	4.5729E-01	1.4559E+02
18b_2016	4.5996E+01	3.2649E+00	1.2708E-02	7.7950E-04	2.2993E-01	3.8700E-06	5.6500E-07	1.4937E-02	1.8852E-02	5.9400E-07	3.1753E-01	1.7965E-01	1.1455E-01	1.5761E+02
wb1a_2015	4.8619E+01	1.2716E+00	2.1586E-03	1.3721E-03	3.2276E-01	2.2800E-05	4.6000E-07	2.3715E-02	4.4758E-03	9.6400E-08	2.0157E-01	8.1020E-02	2.8426E-01	1.0159E+02
wb1b_2015	8.3431E+01	5.4050E+00	1.5642E-02	1.8663E-03	6.6606E-02	1.2900E-05	3.9200E-06	2.4921E-02	1.3302E-02	5.2800E-07	2.0483E-01	4.6575E-02	3.5601E-01	5.2436E+01
wb1c_2015	1.2033E+01	5.2726E+00	1.4565E-03	1.5405E-03	5.2304E-01	1.5800E-05	3.6900E-06	3.0706E-03	1.5156E-03	4.0800E-07	2.2080E-01	5.7715E-02	3.7736E-01	3.0212E+01
wb2a_2015	8.2449E+01	3.1141E+00	4.5287E-03	1.4348E-03	4.3072E-01	1.8900E-05	2.8700E-06	1.9323E-02	4.0739E-03	3.2400E-05	1.4172E-01	3.3852E-02	9.9471E-02	6.3914E+01
wb2b_2015	1.0551E+02	7.8010E+00	3.3825E-03	1.5794E-03	3.0367E-01	5.9600E-05	3.0100E-06	1.7818E-03	2.4130E-02	4.0100E-05	1.3417E-01	1.5506E-01	3.4633E-01	1.6286E+02
wb2c_2015	1.9110E+01	5.5036E+00	1.3783E-02	1.4606E-03	1.8173E-01	2.4600E-05	1.6100E-06	2.2165E-02	3.0118E-02	4.2500E-05	1.3287E-01	1.2961E-01	3.9088E-01	8.3728E+00
wb3a_2015	1.3785E+02	6.7650E-01	1.2288E-02	1.3316E-03	3.9959E-02	4.0300E-05	2.4200E-06	2.9480E-02	1.6036E-03	4.7000E-07	2.0415E-01	2.4059E-01	4.6028E-01	9.5488E+01
wb3b_2015	1.4873E+02	7.6621E-01	1.2102E-02	1.3971E-03	3.1101E-01	4.8200E-05	8.6900E-07	1.6795E-02	2.3779E-02	4.4900E-08	2.2023E-01	2.4705E-01	3.1741E-01	1.2403E+02

wb3c_2015	6.5611E+01	6.2847E+00	1.5157E-02	1.1649E-03	3.4130E-01	1.1700E-05	3.2500E-06	1.5946E-02	1.7366E-02	3.5100E-07	2.3335E-01	2.5768E-01	4.8723E-02	2.2141E+02
wb4a_2015	1.4688E+01	4.5345E+00	1.2809E-02	3.1396E-03	2.6447E-02	5.5100E-05	2.0100E-06	3.0193E-02	1.9883E-02	1.6500E-05	8.0948E-02	7.1451E-02	3.3574E-01	3.3489E+02
wb4b_2015	9.8123E+01	2.5193E+00	2.0119E-03	3.0568E-03	3.7808E-01	4.6700E-05	3.6600E-07	1.4351E-02	3.3961E-03	1.7100E-05	7.3451E-02	3.3393E-02	2.1595E-01	3.7116E+02
wb4c_2015	1.1595E+02	1.6051E+00	9.2700E-05	9.8170E-04	4.7569E-01	5.8400E-05	3.4900E-06	2.2003E-02	7.5211E-03	4.2000E-08	1.1453E-01	1.1880E-01	2.0781E-01	3.9678E+02
wb5a_2015	2.3367E+01	4.3454E+00	9.4660E-04	2.5567E-03	2.4420E-01	1.8800E-05	1.1700E-06	9.0842E-03	2.5312E-02	2.0000E-05	7.6787E-02	3.0580E-02	1.2627E-01	6.6779E+01
wb5b_2015	1.4027E+02	2.3462E+00	1.8647E-03	2.4653E-03	3.8503E-01	2.9600E-05	2.0700E-07	1.4998E-02	1.7163E-02	1.9200E-07	6.5430E-02	6.6677E-02	4.5787E-01	1.4866E+01
wb5c_2015	6.0195E+01	4.6700E-01	4.5043E-03	5.9910E-04	6.3585E-02	5.9800E-05	2.2300E-06	3.4890E-03	2.6172E-02	2.0000E-05	1.0772E-01	1.2659E-01	2.1273E-01	1.2215E+02
wgt1_2015	9.1693E+01	3.9919E+00	7.4969E-03	1.6553E-03	4.4809E-01	1.0100E-05	2.2600E-06	3.2183E-02	2.6563E-02	3.3900E-08	6.1063E-01	1.0484E+00	2.0624E-01	5.3895E+01
wgt2_2015	1.1742E+02	7.2347E-01	3.6169E-03	1.6107E-03	2.0244E-02	2.3100E-05	3.9100E-07	8.5800E-05	5.4406E-03	4.5400E-07	2.4754E-01	3.2122E-01	5.1080E-02	5.7751E+02
wgt3_2015	8.8695E+01	1.1091E+00	7.5226E-03	4.4360E-04	3.4619E-01	4.3000E-05	1.5400E-06	1.2560E-02	2.9139E-02	7.4400E-07	2.5496E-01	3.9016E-01	2.3158E-01	5.3898E+02
wgt4_2015	1.1320E+02	6.4000E+00	5.3149E-03	2.3745E-03	5.4674E-01	3.6400E-05	5.5600E-07	2.0356E-02	6.0460E-03	1.5000E-05	3.1887E-01	5.9451E-01	3.6476E-01	4.8591E+02
wgt5_2015	8.1774E+01	6.7815E+00	8.2642E-03	1.7386E-03	3.1805E-01	1.0100E-05	1.5400E-06	2.3838E-03	1.8379E-02	4.1900E-07	2.8017E-01	3.1226E-01	4.0764E-01	3.3351E+02

3-hydroxybutanone	acetovanillone	alpha-ionone	alpha-terpineol	benzaldehyde	phenyl methanol	phenylmethanethiol	beta-damascenone	beta-ionone	2-phenylethan-1-ol	butyl acetate	butanoic acid	(Z)-3-hexenol	c-whiskylactone	citronellol	decanoic acid	diacetyl
1.0411E+02	7.0568E-01	3.3472E-04	2.4060E-01	6.6878E-01	1.6440E+02	4.9000E-06	8.1939E-04	2.8800E-05	2.5576E+01	8.3652E-01	7.3881E-01	2.9388E-01	3.1839E-02	9.0259E-02	2.2320E+00	6.8201E-02
2.4310E+01	6.3045E-01	1.4447E-03	1.3585E-01	1.7175E+00	1.5021E+02	6.4000E-06	3.5115E-03	7.7000E-05	2.1879E+01	1.2958E+00	7.4797E-01	3.1115E-02	4.0078E-02	9.0097E-02	2.8100E+00	2.4118E-02
5.0055E+01	8.8859E-01	1.2376E-03	1.6492E-01	2.3782E-01	1.3606E+02	5.0000E-06	2.9406E-03	1.2000E-05	2.1504E+01	3.1475E-01	6.3047E-01	1.0322E-02	6.1823E-02	6.4563E-02	2.5562E+00	3.3853E-02
3.0410E+01	7.9523E-01	1.2538E-03	7.7682E-02	1.1670E+00	1.3502E+02	4.9000E-06	6.3686E-04	3.5700E-05	2.3444E+01	1.2349E+00	1.0750E+00	1.9833E-01	5.6753E-02	5.1467E-02	1.7688E+00	9.1185E-02
5.6369E+01	9.4972E-01	2.5484E-03	2.2366E-01	2.5488E-01	1.6742E+00	6.8000E-06	2.0928E-03	8.4100E-06	2.4453E+01	2.7360E-01	1.1038E+00	1.5301E-01	2.5206E-02	6.2827E-02	3.8427E+00	8.7345E-02
1.7863E+01	8.2637E-01	2.3464E-03	8.3886E-02	3.5483E-01	1.2542E+02	4.6000E-06	6.2883E-04	4.5500E-05	2.0609E+01	1.7904E+00	9.2849E-01	3.2365E-01	3.7574E-02	8.4286E-02	1.8712E+00	5.1594E-02
6.3993E+01	9.9187E-01	1.9231E-03	1.8885E-01	1.7071E+00	1.9817E+02	5.2000E-06	2.5361E-03	2.9100E-05	9.9081E+00	4.2530E-01	1.0112E+00	1.5263E-01	2.4858E-02	5.4252E-02	2.7305E+00	4.7570E-02
8.2356E+01	5.6008E-01	1.9625E-03	1.9336E-01	6.0002E-01	1.4451E+02	5.6000E-06	8.8945E-04	8.9600E-05	1.0531E-01	6.9279E-01	1.0264E+00	2.2743E-01	5.8956E-02	9.4155E-02	2.2982E+00	9.3421E-02
2.2633E+01	6.2642E-01	1.8272E-04	1.1607E-01	1.6676E+00	1.4605E+02	5.7000E-06	8.7638E-04	1.1200E-05	1.1736E+01	1.1272E+00	9.2855E-01	3.7235E-01	2.9937E-02	9.5170E-03	1.9798E+00	6.4072E-02
1.2938E+02	9.9227E-01	1.2373E-03	1.4517E-01	1.7375E+00	5.1381E+01	5.5000E-06	2.2053E-03	6.9500E-05	7.1472E+00	1.4595E+00	6.6741E-01	3.9256E-01	3.5171E-02	5.2188E-02	2.5793E+00	3.3172E-02
1.2054E+02	5.2771E-01	2.5335E-03	5.9320E-02	9.0264E-01	1.0497E+02	6.4000E-06	2.3876E-03	5.4300E-05	1.6701E+01	2.3087E-03	8.5188E-01	3.5290E-02	6.1913E-02	4.0240E-02	2.7176E+00	3.2212E-02
4.7733E+01	5.8928E-01	3.9201E-04	3.6153E-03	1.8163E+00	9.9570E+01	4.3000E-06	2.2810E-03	6.0900E-05	1.4594E+01	2.5306E-01	6.9386E-01	2.3284E-01	1.7067E-02	3.3707E-02	2.5501E+00	8.5803E-02
4.0057E+01	1.1786E-01	1.7880E-03	1.2646E-01	1.1082E+00	1.9705E+02	1.2900E-05	1.7008E-03	5.8400E-05	2.4280E+01	1.2924E+00	7.4861E-01	3.9613E-01	4.0077E-02	1.7970E-02	2.5876E+00	6.7351E-02
1.4525E+02	5.3338E-01	1.1444E-03	2.2602E-01	1.3781E+00	9.2700E+01	1.8800E-05	2.4064E-03	5.1300E-05	2.2543E+01	1.0528E+00	7.8354E-01	3.6598E-02	1.7325E-02	7.2880E-02	1.9395E+00	1.0575E-02
8.2345E+01	6.1773E-01	4.9700E-05	1.0597E-01	1.7592E+00	1.5356E+02	1.0700E-05	6.4209E-04	7.6700E-05	2.6841E+01	4.5265E-01	8.2869E-01	3.0786E-01	5.5008E-02	2.5779E-02	2.3178E+00	5.3746E-02
2.9165E+01	3.3979E-01	2.3115E-03	1.4956E-01	1.4401E-01	1.7086E+00	4.1000E-06	1.0198E-03	4.1100E-05	3.0755E+01	3.4691E-01	5.9352E-01	3.6705E-01	4.6344E-02	5.8651E-02	1.4607E+00	8.1472E-03
7.3383E+01	9.0341E-01	2.2114E-03	1.5503E-01	1.1144E+00	9.7102E+01	4.7000E-06	2.1813E-03	1.9600E-05	2.8808E+01	2.5143E-01	5.2480E-01	2.6286E-01	2.9420E-02	1.6530E-03	1.7007E+00	1.5250E-02
1.1856E+02	5.7393E-01	8.1736E-04	1.7296E-01	9.2518E-02	5.8995E+01	3.5000E-06	1.8699E-03	6.3600E-05	2.1998E+01	1.6332E+00	7.5688E-01	3.4701E-01	5.0846E-02	8.3574E-02	1.7802E+00	3.1414E-02
1.0085E+02	1.9941E-01	8.5193E-04	8.8188E-02	1.9009E+00	5.0775E+00	3.9000E-06	1.6121E-03	3.0900E-05	2.4302E+01	1.5256E+00	7.4070E-01	1.0323E-02	3.4256E-02	7.4226E-02	1.9784E+00	3.0193E-02
1.2208E+02	8.3850E-01	1.3526E-03	1.2687E-01	1.6379E+00	1.3915E+02	2.6000E-06	1.0918E-03	8.3000E-05	2.2227E+01	6.9163E-01	8.8117E-01	1.1967E-02	4.9172E-02	6.5775E-02	1.7684E+00	2.5737E-02
5.9178E+01	8.5087E-01	1.7857E-03	1.9913E-01	1.2532E+00	1.1434E+02	2.5000E-06	5.3266E-03	5.1500E-05	1.0731E+01	1.4825E+00	1.3875E+00	3.3730E-02	4.4434E-02	6.9932E-02	2.0488E+00	3.7777E-03
1.3679E+02	6.7110E-01	1.0162E-03	1.6836E-01	1.2260E+00	6.9132E+01	6.1000E-06	1.6585E-03	8.3000E-06	1.4471E+01	1.3352E+00	1.2375E+00	1.0320E-02	6.6501E-02	5.8310E-02	1.9409E+00	4.7553E-02
1.3823E+02	9.9626E-01	1.9325E-03	2.2623E-01	1.2862E-01	1.2042E+02	2.4000E-06	3.3060E-03	8.7900E-05	5.2641E+00	4.7066E-01	1.1618E+00	2.8716E-03	5.3149E-02	1.7693E-02	2.2423E+00	8.4207E-02
4.1373E+01	2.9800E-01	2.0079E-04	1.2057E-01	1.6832E+00	8.7872E+01	2.0300E-05	4.0961E-03	1.8800E-06	1.5318E+01	1.3754E+00	1.7013E+00	7.4206E-02	1.5725E-04	8.7475E-02	2.7169E+00	2.8481E-03
7.2979E+01	8.6792E-01	1.0639E-03	2.2197E-01	1.7222E+00	9.5856E+01	6.0000E-06	2.3387E-03	2.0400E-05	2.8819E+01	1.5446E-01	7.1728E-01	2.8915E-01	3.2405E-02	8.0392E-03	2.4417E+00	4.4120E-02
8.5370E+01	1.7861E-02	2.5326E-03	7.0096E-02	3.3657E-01	1.1260E+02	5.3000E-06	1.5827E-03	4.2600E-05	3.2558E+01	2.8744E-01	8.0850E-01	2.7861E-01	2.1945E-02	1.3680E-02	1.5910E+00	2.7096E-02
1.3422E+02	8.2229E-01	4.5500E-05	1.5503E-01	1.6779E+00	6.7250E+00	2.9000E-06	2.8294E-03	8.3700E-05	4.2460E+01	7.2799E-01	6.0898E-01	8.0042E-02	3.0474E-02	8.2946E-02	1.1270E+00	2.7663E-03
7.0739E+01	7.6312E-01	9.3912E-04	9.2068E-02	1.0490E+00	1.0607E+02	2.2000E-06	1.9972E-03	3.3900E-05	2.1443E+01	8.9370E-01	1.1692E+00	2.4499E-01	5.2304E-02	6.3468E-02	2.1496E+00	7.0979E-03
9.8214E+00	5.9474E-01	3.4103E-04	5.9817E-02	1.4240E+00	8.1827E+01	4.9000E-06	4.7582E-03	1.7400E-05	2.3044E+01	8.9231E-01	1.0019E+00	1.0216E-02	7.5024E-03	2.7438E-02	2.0197E+00	7.7162E-02
6.7981E+01	3.4683E-01	2.5684E-04	8.1176E-03	4.9782E-01	1.0640E+02	5.4000E-06	1.7881E-03	1.1200E-05	1.7358E+01	1.6058E+00	6.8301E-01	2.8522E-01	1.4111E-02	2.1586E-02	2.5245E+00	9.1817E-02
1.2855E+02	1.5035E-01	6.4627E-04	2.2542E-01	9.1987E-01	6.7106E+01	5.6000E-06	6.9344E-04	7.0800E-05	1.9653E+01	1.3388E+00	6.8261E-01	9.1226E-02	5.8084E-02	2.1427E-02	1.2744E+00	9.7692E-03
1.3102E+02	5.3096E-01	2.5970E-03	1.1795E-01	1.5314E+00	1.6028E+02	4.8000E-06	1.1900E-03	6.9300E-05	2.2321E+01	1.4693E+00	5.4455E-01	1.8931E-01	2.8142E-02	2.9659E-02	1.7403E+00	2.5191E-02
1.1652E+02	6.6323E-01	2.2419E-03	4.5656E-02	1.2390E+00	1.7725E+02	7.7000E-06	1.0384E-03	2.1000E-05	4.6069E+01	6.3819E-02	5.5779E-01	8.0039E-02	7.1587E-03	2.6604E-02	1.0017E+00	7.9094E-02
9.0049E+01	6.4429E-01	1.0601E-03	2.0476E-01	1.4613E+00	1.8934E+02	6.0000E-06	3.0256E-03	1.3500E-05	5.5459E+01	1.1475E+00	5.9199E-01	1.2561E-01	2.4056E-02	3.7012E-02	1.1643E+00	3.5791E-02
1.4444E+02	3.6117E-01	1.4042E-03	5.9895E-02	1.8912E+00	1.1247E+02	5.8000E-06	1.3681E-03	6.1400E-05	5.0284E+01	9.9072E-01	6.6201E-01	1.0755E-01	2.3161E-02	2.3467E-02	1.3920E+00	4.9472E-02
1.3718E+02	9.4857E-01	2.5015E-03	5.0012E-02	1.7783E+00	4.1921E+01	5.0000E-06	3.5882E-03	8.1500E-05	5.5323E+01	4.8396E-01	6.5637E-01	9.9661E-02	4.8163E-02	3.5309E-02	1.4973E+00	2.2630E-02
2.5587E+01	6.0943E-01	5.2832E-04	1.4530E-01	2.4908E-01	2.3475E+01	1.3200E-05	7.0274E-03	2.5129E-04	1.6161E+01	9.6952E-01	6.7271E-01	3.4306E-01	5.9306E-02	5.3361E-02	2.0410E+00	2.3872E-02
2.5573E+01	6.6488E-01	2.5789E-03	2.3398E-01	1.6000E+00	4.2674E+01	5.7000E-06	7.9478E-03	3.0079E-04	1.4748E+01	1.4170E+00	7.0561E-01	3.3756E-01	3.6240E-02	1.3507E-02	1.9597E+00	4.5845E-02
5.3041E+00	2.9516E-01	1.9353E-03	5.4653E-02	7.9200E-01	7.2867E+01	2.1800E-05	9.7227E-03	2.6595E-04	1.4335E+01	1.4304E+00	6.9696E-01	3.3956E-01	8.4611E-03	5.5236E-02	3.4979E+00	8.4556E-02
7.2128E+01	3.7954E-01	7.6908E-04	1.9024E-01	1.4681E+00	9.3599E+01	1.2700E-05	6.5174E-03	2.9296E-04	2.1293E+01	8.1349E-01	4.9113E-01	4.6595E-02	6.5911E-02	7.9399E-03	2.5244E+00	9.1778E-02
5.7423E+01	6.8238E-01	8.4696E-04	2.9063E-02	1.3234E+00	1.1481E+02	2.0500E-05	6.1446E-03	2.1570E-04	1.8062E+01	9.8182E-01	4.6920E-01	1.7920E-01	3.9950E-02	1.9191E-02	2.3924E+00	5.6321E-02
1.4271E+02	9.6653E-01	6.6074E-04	2.3201E-01	1.5969E+00	1.0531E+02	1.6200E-05	5.5152E-03	3.3556E-04	1.9010E+01	4.5904E-01	4.7779E-01	2.9641E-01	2.7752E-02	8.7845E-02	2.1643E+00	1.2424E-02
3.8765E+01	6.5546E-01	2.1652E-03	5.2846E-03	8.8850E-01	3.7777E+01	6.0000E-06	8.8346E-03	1.7300E-05	1.6151E+01	1.1735E+00	4.9814E-01	3.2274E-01	3.6166E-03	1.6344E-02	3.0297E+00	6.7675E-02
4.9063E+01	6.1033E-01	1.5090E-03	1.4496E-01	5.2550E-01	2.6472E+01	1.3900E-05	7.9675E-03	1.5500E-05	1.5022E+01	9.2030E-01	5.6065E-01	1.4840E-01	5.9761E-02	7.1400E-03	2.7094E+00	1.2499E-02

3.8111E+01	9.1981E-01	2.0590E-03	6.2505E-02	1.7481E-01	6.6141E+01	7.6000E-06	8.2685E-03	3.9000E-05	1.6126E+01	1.0249E-01	5.3199E-01	1.1328E-01	5.3102E-02	4.2695E-03	2.8134E+00	7.6272E-02
9.4622E+01	8.3021E-01	1.9936E-03	6.7937E-02	4.4310E-01	1.1244E+02	6.2000E-06	5.3251E-03	2.2590E-04	7.7240E+00	4.2943E-01	6.8963E-01	2.6026E-01	1.0977E-02	7.5979E-02	1.9294E+00	2.4144E-02
2.2754E+01	3.3777E-01	5.8966E-04	5.0910E-02	1.1455E+00	9.5744E+01	5.4000E-06	5.8308E-03	2.7207E-04	5.8708E+00	7.3649E-01	6.7056E-01	3.9159E-01	5.4127E-02	4.4520E-02	1.6217E+00	5.2626E-02
9.7018E+01	7.6427E-01	1.3509E-04	1.3754E-01	1.6147E+00	1.6108E+02	1.4000E-06	8.0218E-03	3.5182E-04	3.9884E+00	5.2325E-01	6.8539E-01	9.7436E-02	4.0999E-02	3.6151E-02	2.5615E+00	8.5054E-02
3.8822E+01	7.6370E-01	1.4600E-05	2.4291E-01	1.1996E+00	8.6341E+01	2.9000E-06	7.6371E-03	3.8466E-04	2.7274E+00	2.4138E-01	7.5501E-01	7.3846E-02	7.1167E-03	9.6428E-02	2.3949E+00	7.7614E-02
1.3042E+02	1.6992E-01	2.3984E-03	2.4121E-01	4.5787E-01	6.7948E+01	1.5000E-06	5.4103E-03	2.9851E-04	6.9575E+00	9.4902E-01	7.9931E-01	1.6163E-01	2.3605E-02	5.9538E-02	1.6524E+00	6.6939E-02
2.4203E+01	3.4977E-01	1.9022E-03	2.0167E-01	1.9372E+00	5.3340E+01	3.0000E-07	1.0093E-02	3.4894E-04	1.5150E+01	1.7112E+00	7.8400E-01	1.1406E-01	5.1008E-02	5.9295E-02	1.9782E+00	4.6515E-02
8.2527E+01	1.7116E-01	1.0391E-03	1.6176E-01	1.8932E+00	7.3021E+01	1.4000E-06	9.4070E-03	8.4600E-05	3.5652E+00	1.6689E+00	1.0694E+00	3.0152E-01	6.5255E-02	1.8883E-02	3.3217E+00	4.2429E-02
8.1980E-01	4.2611E-01	4.6923E-04	1.3300E-01	6.5638E-01	1.2598E+02	7.6000E-06	6.7050E-03	5.0900E-06	4.7968E+01	1.6999E+00	7.4069E-01	1.5619E-01	2.4729E-02	9.4328E-02	2.2715E+00	4.3683E-02
5.2840E+01	3.6437E-01	5.6319E-04	6.7728E-02	1.3298E+00	1.1595E+01	8.0000E-07	1.0911E-02	3.3343E-04	5.1090E+00	1.6622E+00	1.2220E+00	2.9335E-01	5.8507E-02	9.2386E-02	2.1467E+00	1.7889E-02
4.1919E+01	4.4885E-01	2.0010E-03	2.3869E-01	1.4393E+00	8.4347E+01	2.9400E-07	7.8489E-03	2.5289E-04	1.2116E+01	4.4214E-01	9.2096E-01	1.6792E-01	6.1116E-02	7.0007E-03	2.9092E+00	6.9464E-02
8.1127E+01	4.9682E-02	1.9506E-03	2.2631E-01	1.6247E+00	7.1329E+01	7.2000E-06	8.6387E-03	2.7846E-04	1.5801E+01	2.4826E-01	1.3644E+00	5.4183E-02	5.2575E-02	8.0109E-02	2.1517E+00	2.5447E-02

diethyl succinate	trans-isoeugenol	ethyl 2-methylbutanoate	ethyl acetate	ethyl butyrate	ethyl cinnamate	ethyl decanoate	ethyl dihydrocinnamate	ethyl hexanoate	ethyl 2-methylpropanoate	ethyl 3-methylbutanoate	ethyl lactate	ethyl octanoate	ethyl propanoate
1.3684E+02	2.1331E-03	6.7105E-03	2.4781E+01	2.4864E-02	5.4060E-04	2.8219E-02	1.2411E-03	6.5213E-01	2.4800E-02	4.3773E-03	3.7224E+00	1.0468E+00	2.8735E+00
5.7603E+01	5.4668E-03	1.4937E-02	3.0297E+01	1.0782E-02	1.9856E-04	1.2605E-01	2.2800E-05	6.8973E-01	2.3050E-02	1.4840E-03	7.2491E+01	8.5227E-01	1.0397E+00
1.9881E+02	4.2964E-03	8.0217E-03	2.6436E+01	1.0992E-02	7.2400E-05	1.1285E-01	1.2528E-03	6.2253E-01	2.8239E-02	3.5689E-04	5.5190E+01	7.8928E-01	2.8269E+00
1.3086E+02	3.7994E-03	1.7433E-02	3.0523E+01	1.9969E-01	1.9286E-04	1.0577E-02	1.0799E-03	8.2595E-01	1.5838E-02	2.1691E-03	1.0720E+02	1.0589E+00	3.1067E+00
1.4176E+01	5.3130E-03	1.3535E-02	3.1015E+01	3.4808E-02	2.3307E-04	1.2513E-01	1.5571E-03	3.4612E-02	1.0134E-02	1.1949E-03	4.2274E+01	1.1948E+00	5.3217E+00
1.6160E+02	1.1048E-03	2.9333E-03	8.9081E+00	2.0886E-01	7.4200E-05	1.4549E-01	6.2599E-04	9.4947E-01	1.8948E-02	2.5222E-03	1.0158E+02	1.1943E+00	4.2421E+00
1.5394E+02	2.5595E-03	4.2197E-03	4.6778E+01	1.0486E-01	2.6046E-04	1.0971E-01	7.0529E-04	1.0776E+00	2.2297E-02	1.3416E-03	1.0797E+02	1.0381E+00	7.0633E-01
2.8626E+01	1.4927E-03	1.4578E-03	4.1894E+01	1.9005E-01	1.6700E-06	1.3554E-01	9.1990E-04	1.0488E+00	2.2347E-03	1.6340E-03	1.3929E+02	1.2112E+00	3.6703E+00
5.8980E+01	2.7697E-03	1.2621E-03	4.2796E+01	2.2853E-01	6.6252E-04	1.2568E-01	1.4405E-03	9.0846E-01	1.3007E-02	2.2719E-04	2.2321E+00	1.0315E+00	8.3364E-02
1.0580E+02	5.3019E-03	4.4433E-03	4.3578E+01	6.9939E-02	9.8911E-04	2.5441E-03	8.5326E-04	8.8009E-01	1.6786E-02	2.5360E-03	9.5657E+00	1.3118E+00	5.2098E+00
6.8689E+01	4.1307E-04	6.3045E-03	3.9808E+01	9.1177E-02	1.0763E-03	1.3259E-01	7.4600E-05	5.9469E-01	1.6623E-02	3.6940E-04	1.2848E+02	7.7887E-01	4.9903E+00
7.8853E+01	5.6440E-03	1.5935E-02	4.1255E+01	1.3747E-01	7.3432E-04	6.1413E-02	5.8052E-04	7.6100E-01	1.6810E-02	1.8532E-03	5.4719E+01	8.2978E-01	2.8637E+00
1.7148E+02	2.9613E-04	9.1126E-03	3.7818E+01	6.5092E-02	2.9775E-04	3.9645E-02	4.4972E-04	1.6215E-02	1.2704E-02	1.3201E-03	1.0271E+01	8.4621E-01	2.8251E-02
6.4824E+01	5.6319E-03	4.6101E-03	4.2398E+01	8.3048E-02	9.9866E-04	1.8239E-01	8.9453E-04	6.4419E-01	2.5596E-02	2.8261E-03	1.2421E+02	7.5038E-01	1.8444E-01
7.3624E+01	3.8928E-03	9.7761E-03	3.2816E+01	1.6991E-02	8.2407E-04	2.3894E-01	3.0970E-04	8.7660E-01	2.0653E-02	1.2624E-03	5.5422E+01	1.8923E+00	2.4680E+00
8.0859E+01	2.6376E-03	6.4358E-03	2.6619E+01	2.6620E-02	9.9080E-04	2.8915E-02	1.1976E-03	4.7602E-01	3.6470E-02	6.1122E-03	1.4737E+02	5.8564E-01	8.3289E-01
1.6921E+02	2.2455E-03	1.4392E-02	2.7269E+01	1.2350E-01	3.0455E-04	9.0698E-02	5.7000E-05	4.3283E-01	3.1204E-02	5.2682E-03	3.8559E+01	4.4859E-02	2.3618E+00
6.5163E+01	5.0562E-03	1.5941E-02	3.2867E+01	1.6509E-01	1.6914E-04	4.0039E-02	8.1000E-05	5.9515E-01	2.4112E-02	3.0533E-03	5.4344E+01	6.0221E-01	8.1807E-01
1.7050E+02	3.7562E-03	1.8197E-03	2.8751E+01	1.1915E-01	3.4189E-04	1.7729E-01	1.1586E-03	5.3500E-01	3.3420E-02	3.8186E-03	6.9426E+01	7.4987E-01	2.2949E-01
1.1284E+02	5.5706E-04	7.7031E-03	3.8079E+01	1.5282E-01	7.5201E-04	1.2796E-01	3.0529E-04	6.5481E-01	2.5348E-02	4.6705E-03	1.4920E+02	8.4530E-01	4.8321E+00
1.8861E+02	2.7184E-03	3.7360E-03	8.4319E+01	1.6329E-01	9.2945E-04	1.4670E-01	2.1700E-05	2.1183E-02	1.6989E-02	1.7451E-03	7.2235E+01	1.0660E+00	5.4933E+00
8.5776E+00	2.5428E-03	1.7038E-02	4.6066E+01	2.0861E-01	7.9608E-04	1.9827E-01	4.5201E-04	8.9630E-01	1.0636E-02	2.8142E-03	5.6560E+01	1.3866E+00	2.7903E+00
7.0373E+01	2.9955E-03	1.5704E-02	4.6232E+01	1.7228E-01	3.2439E-04	1.2114E-01	1.5181E-03	9.2261E-01	1.3811E-03	9.7530E-04	7.0646E+00	8.7780E-01	4.7328E+00
9.2037E+01	7.8462E-03	4.5568E-03	5.4812E+01	2.4785E-01	7.9400E-05	2.4205E-02	1.2348E-03	1.5677E+00	6.1048E-04	2.7151E-03	2.8210E+01	1.3245E+00	5.0745E+00
2.2556E+01	7.2809E-04	6.2947E-03	2.3257E+01	7.2669E-02	6.9026E-04	1.4953E-01	5.1568E-04	7.8035E-01	3.0937E-02	4.9372E-03	1.0924E+02	9.3881E-01	3.6594E+00
8.7346E+01	3.0737E-03	6.9484E-04	2.3101E+01	1.4624E-01	7.1353E-04	8.1086E-02	8.1370E-04	6.9316E-01	3.1116E-02	8.0759E-03	1.2727E+02	4.6647E-01	6.1226E-02
1.2895E+02	6.0092E-03	1.0040E-02	2.4927E+01	5.3868E-02	2.5680E-04	4.9435E-02	1.5820E-03	3.6824E-01	4.3448E-02	9.0868E-03	1.5137E+02	2.6678E-01	1.8473E+00
1.9865E+02	2.2742E-03	1.3690E-02	4.3531E+01	1.6988E-01	4.3443E-04	1.6256E-01	2.3884E-04	9.3606E-01	2.1317E-02	3.8391E-03	5.0271E+01	9.5188E-01	3.1381E-01
1.7739E+02	3.0637E-03	1.5025E-02	3.8481E+01	1.6815E-01	9.7981E-04	1.2258E-01	1.0668E-03	5.6818E-02	1.7932E-02	4.5814E-03	1.1585E+02	2.4615E-01	3.7336E+00
1.9469E+02	4.6139E-03	4.6877E-03	2.9367E+01	4.1132E-02	5.4600E-06	1.0408E-01	1.5725E-03	7.1328E-01	1.7448E-03	1.3621E-03	1.4863E+02	9.5077E-01	3.3612E+00
1.0552E+02	3.9692E-03	1.4357E-02	2.1269E+01	1.2393E-01	1.0479E-04	1.5513E-02	2.4992E-04	5.3715E-01	1.1275E-02	3.7596E-03	6.6488E+01	6.8750E-01	1.6939E+00
1.8971E+02	1.0017E-03	6.7261E-03	2.7943E+01	7.2448E-03	1.0113E-03	1.1467E-01	1.2633E-03	4.6228E-01	5.2607E-02	1.2896E-03	1.2610E+02	5.4184E-01	1.0676E+00
5.4633E+01	4.3407E-03	8.6280E-03	2.7869E+01	7.0499E-02	8.7141E-04	1.9504E-01	6.5722E-04	9.0209E-01	4.4377E-02	9.8151E-03	7.6861E+01	1.3370E+00	2.4565E+00
2.3403E+01	3.8093E-03	1.1002E-02	2.9688E+01	1.5285E-02	4.2757E-04	3.9621E-02	2.7077E-04	8.6636E-01	4.8124E-02	1.2130E-02	3.5722E+01	1.4441E+00	3.9154E+00
7.5351E+00	1.0976E-03	1.4042E-04	2.3956E+01	6.3812E-02	3.0824E-04	8.5508E-02	1.8298E-03	1.3200E+00	4.7011E-02	1.1078E-02	1.3118E+02	1.5071E+00	3.7299E-01
1.7055E+02	3.0435E-03	1.0185E-02	2.3615E+01	8.9242E-02	7.9226E-04	1.5102E-04	1.8580E-03	1.6944E+00	4.2039E-02	8.0373E-03	6.1040E+01	1.1891E+00	2.2906E-01
1.0742E+02	4.6545E-03	2.9224E-04	6.2324E+01	2.1935E-01	3.5449E-04	3.3238E-02	7.5036E-04	4.9057E-03	4.0100E-02	2.9349E-04	1.3529E+02	2.7271E-01	3.8496E+00
9.2151E+01	1.5579E-03	5.6300E-05	5.2850E+01	2.2202E-01	9.8540E-04	5.0564E-02	1.5153E-03	2.5440E-02	4.3402E-02	1.7601E-03	9.8661E+01	6.2691E-01	7.8089E-01
1.8991E+02	5.0588E-03	8.3196E-03	5.1317E+01	2.1761E-01	4.5252E-04	2.1313E-02	3.1410E-04	6.0334E-02	4.3364E-02	2.6204E-03	1.3002E+01	1.1720E+00	4.7164E+00
4.4126E+01	4.6331E-03	3.9829E-03	5.4464E+01	8.1055E-02	2.9074E-04	9.9421E-02	9.1570E-04	2.9227E-02	4.6272E-02	2.0639E-03	1.2265E+02	7.8503E-01	5.2982E+00
1.6332E+02	5.6471E-03	1.6469E-02	5.9439E+01	1.2914E-01	1.5622E-04	1.6403E-01	5.7591E-04	4.2486E-02	4.2000E-02	1.3758E-03	2.0699E+01	6.4835E-01	4.3946E+00
1.3034E+02	1.8485E-03	2.6952E-03	6.0000E+01	1.5003E-01	1.5483E-04	1.2689E-01	1.4567E-03	9.0184E-03	3.5587E-02	2.9290E-03	5.2905E+01	1.2000E+00	1.2001E+00
1.4905E+02	4.7674E-04	8.2219E-03	4.7631E+01	1.3875E-01	6.3863E-04	6.3001E-02	2.7767E-04	3.0742E-04	5.3914E-02	2.4446E-03	3.3662E+01	1.7937E+00	1.8509E+00
1.6223E+02	3.6679E-03	1.7805E-02	5.0048E+01	1.5893E-01	2.0585E-04	3.4733E-02	6.6144E-04	5.2522E-02	4.3007E-02	2.4989E-03	4.6506E+01	1.5927E+00	4.6405E+00

6.1423E+01	4.6601E-03	1.4256E-02	5.1049E+01	1.5748E-01	9.2600E-05	1.4436E-01	1.4944E-03	5.3990E-02	4.3509E-02	2.8543E-03	5.2021E+01	1.7574E+00	5.0979E+00
1.2671E+02	2.2832E-03	1.6344E-02	8.0779E+01	1.5911E-01	5.0365E-04	3.9480E-02	1.4787E-03	5.7551E-02	5.1318E-02	2.9987E-03	1.1660E+02	2.1582E+00	2.9149E+00
1.7122E+02	8.2672E-04	1.7088E-02	8.2215E+01	1.6284E-01	6.5163E-04	1.8884E-01	1.4680E-03	9.3410E-03	4.2087E-02	2.2058E-03	7.6350E+00	1.6541E+00	1.0235E+00
1.3677E+02	1.6688E-03	2.5484E-03	9.2822E+01	1.6087E-01	4.8963E-04	1.0957E-01	1.3831E-03	3.2267E-02	4.7993E-02	5.5358E-04	4.4400E+01	2.0106E+00	1.9372E+00
6.2035E+00	4.4403E-03	2.9494E-03	6.3773E+01	1.5518E-01	3.0181E-04	9.3864E-03	5.1200E-05	5.7158E-02	4.2223E-02	2.3682E-04	1.2719E+02	2.4595E+00	4.4892E+00
1.6973E+02	7.0300E-05	1.2579E-02	6.7782E+01	1.7310E-01	3.1217E-04	1.8801E-01	8.3953E-04	1.7043E-02	4.3636E-02	1.3022E-03	1.5390E+02	2.0852E+00	3.9805E+00
3.3653E+01	1.4601E-03	2.1438E-03	6.9684E+01	1.8136E-01	1.0567E-04	1.5368E-01	9.1651E-04	3.2469E-03	4.2093E-02	1.1855E-03	5.7534E+01	2.5759E+00	4.7715E+00
1.7774E+02	5.8640E-03	1.0771E-02	5.5883E+01	3.3607E-01	1.9841E-04	6.6389E-03	7.7352E-04	2.6375E-02	1.3147E-03	2.4461E-03	1.3752E+02	2.4971E+00	1.9052E+00
9.0273E-01	2.2677E-03	1.4349E-02	8.7078E+01	1.8062E-01	2.0816E-04	1.8228E-01	1.3268E-03	3.8553E-02	3.8821E-02	2.4126E-03	2.8956E-01	2.1457E+00	5.2492E+00
1.7076E+02	3.1537E-03	5.3575E-03	9.2977E+01	3.2863E-01	8.7612E-04	1.9124E-01	5.9500E-05	2.7105E-02	1.6618E-02	1.5610E-03	6.5943E+01	2.3246E+00	4.3940E+00
1.1367E+02	3.7736E-03	1.5499E-02	7.4983E+01	2.2017E-01	9.7452E-04	8.8708E-02	1.1534E-03	2.0458E-02	5.3469E-02	2.6128E-03	1.9926E+01	3.2551E+00	9.9382E-01
1.8480E+02	5.7938E-03	1.5036E-02	6.2243E+01	2.8720E-01	4.2020E-04	9.9748E-02	6.2349E-04	1.3741E-02	2.4387E-02	2.4825E-03	3.5258E+01	2.7962E+00	4.6189E+00

ethyl vanillate	eugenol	2-furfurylthiol	gamma-decalactone	gamma-nonalactone	oxolan-2-one	geraniol	2-methoxyphenol	hexanoic acid	hexyl acetate	3-methylbutyl acetate	3-methylbutan-1-ol	2-methylpropan-1-ol	2-methylpropylethanoate	2-methylpropanal
7.5894E-01	1.5824E-03	3.6000E-06	2.3153E+01	3.3800E-05	9.5888E-03	5.1669E-03	7.6597E-03	4.0609E+00	1.0960E+00	4.5824E-01	1.2125E+02	3.5487E+00	2.0754E-01	8.1000E-03
9.5353E-01	2.7971E-03	4.3000E-06	2.0086E+01	3.8636E-03	2.4168E-02	8.6302E-03	9.1683E-03	3.7172E+00	9.4218E-01	5.4842E-01	1.1959E+02	3.7837E+01	9.9820E-03	7.8000E-03
5.7925E-01	4.5134E-03	4.6000E-06	1.3139E+01	3.0068E-03	2.2832E-02	8.3882E-03	4.1462E-03	3.3277E+00	9.3161E-01	4.6441E-01	1.1589E+02	1.3480E+01	1.4272E+00	9.4000E-03
9.5663E-01	5.9800E-05	2.4000E-06	9.7599E+00	1.2953E-04	1.0222E-02	3.2769E-03	1.6312E-03	5.8168E+00	1.4949E+00	9.8491E-01	1.3174E+02	1.1359E+00	1.5364E+00	8.4000E-03
1.5056E-01	6.5056E-04	3.2000E-06	2.2810E+01	4.4900E-05	1.0190E-02	8.1632E-03	1.7641E-03	6.0914E+00	9.0832E-01	1.1803E+00	1.2515E+02	3.1944E+01	7.5693E-01	9.6000E-03
9.1156E-01	1.9447E-03	3.7000E-06	2.8382E+01	3.0887E-04	2.1588E-02	1.6866E-02	4.1571E-03	5.2489E+00	1.6111E-01	1.1970E+00	1.2389E+02	1.1935E+01	1.2684E+00	8.7000E-03
7.4528E-01	3.1653E-03	2.0000E-06	2.5261E+01	4.7916E-04	6.5942E-03	1.2402E-02	5.5317E-03	4.9726E+00	8.7224E-02	1.5834E+00	1.1112E+02	2.9462E+01	1.1845E+00	1.2100E-02
9.1200E-01	2.3362E-03	1.4000E-06	6.1334E+00	5.7740E-04	2.4814E-02	1.4369E-02	8.5208E-03	5.7274E+00	4.4625E-01	1.5334E+00	1.0767E+02	3.2951E+01	8.4643E-02	1.2000E-02
2.5401E-01	1.2707E-04	1.5000E-06	1.3086E+01	1.7544E-04	2.3687E-02	1.1939E-02	1.5474E-03	5.5906E+00	3.5723E-01	1.4326E+00	1.0326E+02	2.0114E+00	1.5612E+00	1.1800E-02
6.3223E-01	1.7982E-03	2.9000E-06	1.7604E+01	5.1397E-04	2.6302E-03	3.2045E-03	7.5309E-03	3.2529E+00	1.2200E+00	7.7949E-01	1.0314E+02	1.4973E+01	1.0093E+00	1.2900E-02
6.3200E-01	2.9819E-03	3.1000E-06	2.8398E+01	7.7705E-03	5.2596E-03	1.9664E-03	7.9230E-03	4.1785E+00	4.6641E-01	6.6590E-01	1.1168E+02	5.5890E+00	4.6094E-01	1.0600E-02
1.0142E-01	2.3588E-03	2.2000E-06	5.3749E+00	1.5084E-03	1.6281E-02	9.4742E-03	3.3162E-03	3.5151E+00	2.5158E-01	7.4864E-01	1.0916E+02	1.9135E+01	8.5516E-01	1.1000E-02
9.1839E-01	3.0755E-03	2.8000E-06	1.9921E+01	3.3053E-04	1.3611E-02	1.7987E-03	5.8163E-04	4.5982E+00	8.8658E-02	1.0615E+00	9.7870E+01	3.1546E+01	6.3029E-01	1.1700E-02
7.7589E-01	2.6074E-03	1.8000E-06	8.4406E-01	8.3920E-04	2.0636E-02	8.1872E-04	3.5688E-03	4.2245E+00	1.1138E+00	8.6799E-01	1.0053E+02	1.3656E+01	1.1799E+00	1.2600E-02
8.2859E-01	4.4128E-03	2.5000E-06	4.5498E+00	6.5427E-04	1.1928E-02	6.9581E-03	5.0209E-03	4.7827E+00	1.2826E+00	1.0559E+00	9.6635E+01	3.3476E+01	6.8214E-01	9.9000E-03
6.5985E-01	3.8770E-03	4.0000E-06	2.3069E+01	4.5100E-05	9.3093E-03	1.1271E-02	1.4411E-03	2.9092E+00	4.1056E-01	2.5469E-01	1.5043E+02	1.9036E+01	6.4027E-01	8.9000E-03
2.2499E-01	4.8189E-03	4.8000E-06	2.3919E+01	6.5562E-03	8.1279E-03	6.4336E-03	9.3459E-03	2.4892E+00	6.7945E-01	2.6361E-01	1.4910E+02	8.3248E+00	2.5286E-01	8.4000E-03
8.6851E-01	3.6342E-03	3.9000E-06	2.8847E+01	1.0974E-04	2.0764E-02	1.7353E-02	6.5298E-03	2.8843E+00	1.3590E+00	8.2507E-01	1.5344E+02	2.4451E+01	7.5923E-01	8.8000E-03
3.4100E-01	1.4908E-03	4.0000E-06	6.1259E+00	5.1931E-04	6.3544E-03	1.2671E-02	2.6656E-03	3.3046E+00	1.1616E+00	6.4506E-01	1.4800E+02	1.1589E+01	4.6079E-02	8.8000E-03
7.1978E-02	3.4815E-03	3.3000E-06	2.7785E+01	5.2261E-04	5.5280E-03	1.6036E-02	9.0605E-03	3.6907E+00	3.6406E-01	9.9020E-01	1.4625E+02	1.4015E+01	8.3165E-01	9.3000E-03
6.8070E-01	1.2005E-03	1.5000E-06	1.2218E+01	3.0971E-03	1.4156E-02	3.2472E-03	2.9796E-03	6.1098E+00	8.5994E-01	1.4582E+00	8.8936E+01	2.1615E+01	9.4663E-01	8.8000E-03
2.9212E-01	5.5216E-03	1.4000E-06	3.2568E+01	2.5856E-04	2.0756E-03	2.3765E-03	7.5422E-03	6.7338E+00	3.1641E-01	1.7759E+00	1.1820E+02	3.2236E+01	4.9542E-01	1.1300E-02
3.3266E-01	5.6810E-03	1.7000E-06	2.1484E+01	3.7100E-05	1.6641E-02	1.5089E-02	2.8206E-03	5.3113E+00	3.9001E-01	1.3210E+00	1.3111E+02	1.6282E+01	1.0648E+00	1.0200E-02
8.3942E-01	4.2992E-04	2.2000E-06	2.4830E+01	6.0906E-04	2.0590E-02	1.2942E-02	1.8777E-03	8.1130E+00	7.7019E-01	2.6118E+00	1.0028E+02	1.5497E+01	1.5968E+00	9.4000E-03
3.8904E-02	4.9405E-03	3.2000E-06	2.8455E+01	3.0261E-04	2.1461E-03	1.4725E-02	7.2549E-03	4.2522E+00	9.5979E-01	6.7687E-01	1.8473E+02	3.1360E+00	1.2520E+00	1.0900E-02
6.8860E-01	2.8672E-03	1.6000E-06	2.7901E+00	4.5900E-05	1.7055E-02	1.0899E-02	5.0489E-03	4.8568E+00	1.3622E+00	6.7903E-01	1.7360E+02	3.8830E+01	8.4010E-01	8.9000E-03
1.5504E-01	1.9394E-03	1.7000E-06	7.6618E+00	1.3565E-03	2.1058E-02	1.1096E-02	6.3929E-03	2.9233E+00	1.0361E-01	4.5628E-01	2.5018E+02	1.2635E+01	9.6964E-01	1.2100E-02
6.0838E-01	5.3486E-03	2.1000E-06	3.9172E+00	6.9667E-03	4.4539E-03	4.6125E-03	2.6756E-03	5.3424E+00	7.8639E-01	1.4483E+00	1.6821E+02	7.2646E+00	3.7838E-01	8.5000E-03
5.1277E-01	8.4813E-04	4.9000E-06	2.1872E+01	1.7703E-04	1.2353E-02	1.5791E-02	1.0437E-02	5.0359E+00	3.7782E-01	1.4513E+00	1.5918E+02	3.5923E+01	6.4487E-01	1.1100E-02
1.5821E-01	2.9445E-03	3.4000E-06	2.2316E+01	2.1082E-04	1.4954E-02	4.8012E-03	4.5614E-03	3.6254E+00	1.0419E+00	4.0458E-01	1.0135E+02	5.7991E+00	7.3439E-02	7.0000E-03
3.3848E-01	2.8878E-03	1.7000E-06	5.5896E+00	2.5377E-04	2.2184E-02	1.0104E-02	9.4675E-03	3.9360E+00	1.0365E+00	3.5094E-01	1.2142E+02	1.5664E+01	1.9874E-01	1.2000E-02
5.2661E-01	1.5319E-03	2.4000E-06	5.3143E+00	2.3958E-04	7.3752E-03	1.5266E-02	1.6263E-03	3.6644E+00	7.8002E-01	4.8607E-01	1.5128E+02	4.6397E+01	7.0368E-01	1.1000E-02
7.6337E-01	1.9028E-03	3.3000E-06	2.4432E+01	6.7274E-04	4.9722E-03	1.5774E-02	3.7055E-03	1.9504E+00	1.1400E+00	5.6331E-01	2.0376E+02	2.2089E+01	1.0138E+00	1.3000E-02
4.2535E-01	4.3172E-03	3.3000E-06	2.3397E+01	3.4570E-04	1.3111E-02	1.3110E-02	8.6268E-03	1.9826E+00	2.2854E-01	6.0085E-01	2.2043E+02	1.7935E+01	1.4621E+00	1.4000E-02
6.2137E-01	1.8750E-03	1.7000E-06	2.8103E+01	1.2585E-04	7.2537E-03	2.2789E-03	6.2623E-03	2.1434E+00	7.6824E-01	2.1513E-01	1.7332E+02	1.2131E+01	1.0720E+00	1.2900E-02
5.6197E-01	4.1913E-03	2.2000E-06	3.1690E+01	1.8993E-04	1.9108E-02	9.6677E-03	6.7928E-03	1.8452E+00	7.3785E-01	2.3142E-01	1.6799E+02	2.0673E+00	1.1101E+00	1.3000E-02
2.3744E-02	1.1051E-04	2.9700E-07	2.2905E+01	6.4855E-04	1.2358E-02	1.0579E-02	8.2953E-03	5.7193E+00	4.4290E-01	1.0213E+00	1.3540E+02	3.6901E+01	1.1919E+00	7.8900E-03
3.4860E-01	1.2937E-04	2.3200E-07	3.0350E+00	5.5095E-04	2.1757E-02	1.7540E-02	4.8271E-03	5.2164E+00	8.6571E-01	8.3034E-01	1.4122E+02	1.0831E+01	8.0940E-01	7.5000E-03
5.2932E-01	1.4282E-03	1.6600E-07	2.5211E+01	2.2137E-04	8.4062E-03	2.9132E-03	7.1865E-04	5.3665E+00	1.2942E+00	1.0305E+00	1.3834E+02	2.2433E+01	9.5018E-01	7.7500E-03
9.8762E-01	5.1054E-03	6.1000E-06	3.3345E+01	4.8439E-04	2.0221E-02	1.7981E-02	5.0496E-03	3.5133E+00	1.2322E+00	6.3234E-01	1.3985E+02	2.7100E+01	4.8030E-01	7.3700E-03
2.0961E-01	5.5357E-03	6.1000E-06	2.9226E+01	2.4274E-04	8.9670E-03	4.7977E-04	1.5032E-04	3.4201E+00	3.6532E-01	5.9495E-01	1.2606E+02	1.2172E+01	3.8257E-02	6.2400E-03
6.5818E-01	3.8372E-03	4.8000E-06	8.5824E+00	3.0348E-04	6.5905E-04	2.0365E-03	7.5952E-03	2.7546E+00	9.1573E-01	6.9503E-01	1.2742E+02	1.1390E+01	9.8798E-01	6.5200E-03
8.7185E-01	4.4607E-03	3.8000E-06	1.7399E+01	1.1400E-05	3.8148E-03	2.2640E-03	9.0149E-03	2.9253E+00	1.0097E+00	5.3383E-01	1.0344E+02	1.2976E+01	2.2768E-02	2.1996E-03
2.1781E-01	1.2311E-03	3.3000E-06	7.8808E+00	2.5353E-04	1.4951E-02	1.4584E-02	1.2362E-03	3.2101E+00	4.6072E-01	5.0992E-01	1.0849E+02	9.9232E+00	7.5818E-02	2.8023E-03

5.3374E-01	1.7870E-03	3.0000E-06	2.8986E+01	5.3249E-04	1.6786E-02	1.8684E-02	2.4677E-03	3.0457E+00	5.3591E-01	4.9319E-01	1.0236E+02	2.8819E+01	1.5864E+00	6.0700E-03
2.8542E-01	2.8914E-03	3.8000E-06	3.1452E+01	1.7097E-04	2.1423E-02	4.2027E-03	6.9345E-03	3.6364E+00	4.9832E-01	3.2073E-01	1.3194E+02	2.7132E+01	1.3803E+00	1.1490E-02
4.5358E-02	4.6711E-03	3.4000E-06	1.7121E+01	1.1282E-04	9.9149E-03	3.1505E-03	4.4292E-03	3.5780E+00	1.1493E+00	3.0451E-01	1.2158E+02	1.7566E+01	1.5379E+00	1.0240E-02
3.6962E-02	5.4621E-03	3.0000E-06	2.3977E+01	4.7981E-04	1.4404E-02	3.0917E-03	7.9885E-03	3.7565E+00	5.6609E-01	3.1458E-01	1.2901E+02	2.3465E+00	1.0722E+00	9.5700E-03
1.2803E-01	4.6292E-03	3.6000E-06	2.7246E+01	6.6485E-04	2.2682E-02	4.0606E-03	5.0207E-03	3.6552E+00	1.2888E+00	1.8340E-01	9.6922E+01	1.5003E+01	1.5018E+00	9.7400E-03
6.8902E-01	3.8202E-03	3.9000E-06	3.2106E+01	3.6366E-04	1.9158E-02	1.6481E-02	5.0172E-03	3.1081E+00	1.2642E+00	2.1898E-01	9.3720E+01	1.4142E+01	4.6979E-01	8.5900E-03
4.8217E-01	4.3051E-03	3.5000E-06	7.9525E+00	9.3100E-05	1.6231E-02	1.3899E-02	3.4761E-03	3.2931E+00	1.1914E+00	2.6659E-01	1.1074E+02	3.7069E+01	1.0369E+00	8.1100E-03
9.6093E-01	3.5122E-03	3.0500E-07	2.2172E+01	6.5663E-04	1.1097E-02	6.7315E-03	4.6489E-03	5.4199E+00	1.1506E+00	2.3786E+00	1.2001E+02	3.7005E+01	1.0624E-01	3.7207E-03
4.9969E-01	6.8280E-03	1.1300E-05	1.9976E+01	8.6300E-05	1.0980E-02	4.0002E-03	8.7772E-03	3.5795E+00	2.4092E-01	8.6108E-01	1.1143E+02	2.5443E+01	9.0009E-01	7.1300E-03
1.7093E-01	2.2821E-03	4.8000E-06	1.3966E+01	1.9440E-04	7.1567E-03	1.0430E-02	5.5329E-03	3.8060E+00	1.7407E-01	1.8291E+00	1.3664E+02	9.7180E-01	5.9442E-01	6.4800E-03
4.8203E-01	4.0465E-03	1.4000E-06	2.0067E+01	6.9294E-04	1.8195E-02	1.1838E-03	5.0826E-03	3.6223E+00	1.1624E+00	9.7835E-01	1.1868E+02	3.4759E+01	6.2284E-01	2.9911E-03
6.8452E-01	3.7573E-03	2.5000E-06	2.7392E+01	6.7550E-04	2.1421E-02	1.4542E-02	2.7100E-03	4.4503E+00	2.8690E-01	9.3842E-01	1.2157E+02	3.6692E+01	5.8410E-01	6.8400E-03

isobutyric acid	3-methylbutanal	3-methyl butanoic acid	linalool	3-methylphenol	methional	methyl vanillinate	methionol	2-methylphenol	octanoic acid	2-phenylacetaldehyde	ethyl phenylacetate	trans-whiskylactone	vanillin
1.7105E+00	1.0200E-02	3.8138E-01	8.3561E-02	3.2495E-02	1.1800E-02	3.1987E-01	1.4196E+00	5.8222E-03	1.2235E+01	2.1600E-01	2.3905E-01	5.4677E-01	9.6393E-01
8.2927E-01	9.9000E-03	3.5919E-01	8.6777E-02	1.1768E-02	7.5000E-03	2.5802E+00	1.2750E+00	2.8315E-03	1.1559E+01	2.0340E-01	3.5868E-03	3.7968E-01	2.0747E-01
7.9843E-01	1.0500E-02	3.6731E-01	8.6953E-02	5.4386E-02	1.3400E-02	1.7413E+00	1.3050E+00	2.5973E-02	1.0315E+01	2.0760E-01	1.7731E-01	6.6196E-01	7.4444E-01
1.6914E+00	1.6900E-02	3.0805E-01	7.8298E-02	6.5004E-02	4.6000E-03	2.9813E+00	8.4078E-01	2.7986E-02	1.6374E+01	9.0000E-02	1.9072E-01	7.7969E-01	3.0174E-01
5.5480E-01	1.6100E-02	3.4595E-01	7.6508E-02	1.8938E-02	5.5000E-03	1.1965E+00	3.1763E-01	2.5759E-02	1.8658E+01	1.0300E-01	1.5927E-01	4.5473E-01	3.5971E-01
1.6715E+00	1.6400E-02	3.0666E-01	8.3397E-02	5.5495E-02	5.5000E-03	1.6124E+00	2.5828E-01	2.4183E-02	1.6523E+01	8.7500E-02	2.4697E-01	1.9782E-02	5.9536E-01
7.9573E-01	1.8900E-02	3.1772E-01	7.6382E-02	4.8332E-02	6.3000E-03	9.7690E-01	8.0459E-01	1.4387E-02	1.4135E+01	1.0940E-01	2.1210E-01	4.0191E-01	2.5306E-01
8.7553E-01	1.8700E-02	3.3761E-01	7.0166E-02	2.3605E-02	5.1000E-03	2.2680E+00	8.4940E-01	1.0565E-02	1.6053E+01	1.0810E-01	1.6203E-01	1.7340E-01	1.9752E-01
2.1989E+00	1.8300E-02	2.8264E-01	7.0695E-02	1.3792E-02	4.7000E-03	3.4074E-01	7.0395E-01	1.4263E-02	1.7349E+01	1.1010E-01	1.0072E-01	4.7429E-01	6.5424E-01
4.0450E-02	1.6300E-02	2.3372E-01	5.2854E-02	4.3242E-02	7.1000E-03	5.2850E-01	7.2405E-01	2.1019E-02	1.0333E+01	1.1850E-01	9.4657E-02	7.1923E-02	8.2522E-02
1.2241E+00	1.6700E-02	2.5549E-01	5.6789E-02	5.1051E-02	6.2000E-03	1.5478E+00	6.3417E-01	1.1794E-02	1.3998E+01	1.1790E-01	1.6851E-01	2.5514E-01	6.5160E-01
1.4688E+00	1.6700E-02	2.2969E-01	4.1897E-02	3.5924E-03	6.2000E-03	1.0849E+00	9.7888E-01	1.9650E-02	1.0659E+01	1.2520E-01	2.0032E-01	9.9116E-03	9.4941E-01
3.3544E-01	1.2200E-02	2.3247E-01	2.9488E-02	3.8173E-04	6.3000E-03	2.9019E+00	8.2231E-01	1.9243E-02	1.4466E+01	1.6010E-01	1.3783E-01	3.9309E-01	1.3207E-01
1.2042E+00	1.2600E-02	2.1529E-01	2.9741E-02	8.5990E-03	9.2000E-03	9.6578E-01	3.8131E-01	1.7159E-02	1.4082E+01	1.4760E-01	6.6243E-02	4.1626E-01	3.4118E-01
1.8244E+00	1.3200E-02	2.7748E-01	2.6904E-02	6.7447E-02	7.4000E-03	4.6041E-01	7.8606E-01	6.2514E-04	1.5596E+01	1.4730E-01	6.2388E-02	3.4575E-01	8.1241E-01
1.2980E-01	1.3800E-02	6.6148E-01	9.0929E-02	4.3362E-02	1.5600E-02	1.7043E+00	1.3810E+00	4.9483E-03	8.6429E+00	9.8400E-02	4.8279E-03	7.8004E-01	9.7192E-01
1.7430E+00	1.0500E-02	6.5058E-01	9.1697E-02	5.9741E-02	1.1900E-02	9.1014E-01	1.2309E+00	3.0060E-02	7.8021E+00	8.6000E-02	1.0693E-01	4.8397E-01	1.8544E-01
8.1653E-01	1.3600E-02	5.1967E-01	7.1911E-02	4.5307E-02	1.2200E-02	1.6239E+00	1.5210E+00	7.6358E-03	9.0465E+00	7.9000E-02	1.3787E-01	7.6592E-01	1.1166E-01
1.5625E+00	1.1300E-02	4.6648E-01	7.3202E-02	1.1022E-02	1.0000E-02	2.5767E+00	1.4359E+00	2.4728E-02	9.5047E+00	7.3000E-02	1.0652E-01	5.5489E-01	8.0878E-01
1.7600E+00	1.0600E-02	4.6164E-01	6.5133E-02	1.0262E-02	1.2700E-02	2.4833E+00	1.2660E+00	1.2317E-02	1.1376E+01	7.3900E-02	1.7406E-01	5.0884E-01	1.8414E-01
4.2601E-01	1.2800E-02	2.1338E-01	7.2400E-02	5.7507E-02	4.4000E-03	8.9585E-01	8.4355E-01	2.2435E-02	1.4855E+01	1.5620E-01	1.9525E-02	7.8151E-01	2.4437E-01
1.3025E+00	3.5600E-02	3.5351E-01	9.1984E-02	1.7992E-02	6.5000E-03	1.8727E+00	1.3300E-01	2.0685E-02	1.7355E+01	8.6000E-02	9.1843E-02	4.1096E-01	1.9911E-01
9.0054E-01	1.8300E-02	2.4485E-01	8.5694E-02	1.9163E-02	4.7000E-03	1.0857E+00	1.8686E-01	9.6943E-03	1.3634E+01	1.5500E-01	3.2243E-03	1.1813E-01	3.7258E-01
1.4693E+00	1.8200E-02	3.5391E-01	4.7847E-02	1.2815E-02	5.7000E-03	1.5151E+00	3.9934E-01	2.4429E-02	2.0598E+01	1.6870E-01	9.0886E-02	1.5908E-02	2.8964E-01
2.2746E+00	1.5900E-02	7.6589E-01	1.1463E-01	3.9975E-02	2.1600E-02	2.2635E+00	1.3695E+00	2.2031E-02	1.3035E+01	1.0490E-01	1.9649E-01	6.5456E-01	8.6671E-01
2.1168E+00	1.4900E-02	9.2751E-01	1.1722E-01	1.3763E-02	1.8100E-02	2.3189E+00	1.4798E+00	8.5141E-03	1.3672E+01	1.1780E-01	2.0799E-01	4.0823E-01	3.1680E-01
2.2632E+00	2.1700E-02	1.1016E+00	9.9169E-02	2.1055E-02	2.4100E-02	1.5244E+00	1.7407E+00	2.3176E-02	7.3354E+00	1.4240E-01	3.8184E-02	4.4210E-01	7.0152E-01
3.8327E-01	1.7100E-02	5.8083E-01	7.6562E-02	5.1400E-02	1.7800E-02	2.7064E+00	1.1708E+00	2.8214E-02	1.6278E+01	8.7200E-02	4.6832E-02	1.1163E-02	9.5141E-01
3.5369E-01	1.6400E-02	5.1184E-01	7.4059E-02	5.3618E-02	1.8600E-02	1.1806E+00	1.1923E+00	1.9850E-02	1.5215E+01	8.3400E-02	1.2546E-01	2.4034E-01	2.8810E-01
8.9678E-01	1.2000E-02	2.7768E-01	6.2514E-02	1.3218E-02	9.0000E-03	2.2680E+00	1.0093E+00	1.3453E-02	1.1122E+01	1.9200E-01	3.4034E-02	6.4684E-01	5.4008E-01
2.0210E+00	1.8000E-02	4.5012E-01	8.7184E-02	1.6799E-02	1.1000E-02	2.8608E+00	1.2308E-02	1.5686E-02	1.2484E+01	2.1400E-01	1.1017E-01	4.1600E-02	1.3429E-01
3.0035E-01	1.7000E-02	4.5790E-01	7.3497E-02	2.1229E-02	1.5000E-02	8.8081E-01	1.9614E+00	1.0756E-02	1.0554E+01	2.1500E-01	3.7483E-03	4.9598E-01	9.8349E-01
2.9457E-02	1.6000E-02	1.0369E+00	7.9949E-02	3.6189E-02	1.3000E-02	2.7627E+00	1.5713E+00	2.3428E-02	7.4325E+00	2.2900E-01	2.4113E-01	6.2064E-03	9.5541E-01
9.8788E-01	2.1000E-02	1.0139E+00	7.9207E-02	6.6930E-02	1.2000E-02	2.4661E+00	1.6736E+00	1.5606E-02	7.0837E+00	2.6600E-01	6.1514E-04	7.7652E-01	4.4330E-01
2.2152E+00	2.1000E-02	1.2260E+00	7.9909E-02	5.1350E-02	1.5000E-02	2.6414E+00	1.1446E+00	2.2373E-02	8.9886E+00	2.1900E-01	1.2945E-01	3.8217E-01	5.9809E-02
1.4470E+00	2.0000E-02	1.2231E+00	8.2119E-02	5.1560E-02	1.4000E-02	1.7227E+00	1.0730E+00	3.6021E-03	8.6271E+00	2.6100E-01	9.7604E-02	2.5280E-01	9.1476E-01
1.0321E+00	8.7000E-03	2.4487E-01	6.9131E-02	1.1267E-02	4.0900E-03	1.1021E+00	7.7042E-01	2.8859E-02	1.3734E+01	6.2850E-02	9.3891E-03	2.0816E-01	4.8231E-01
2.2234E+00	7.8300E-03	2.0577E-01	7.6801E-02	2.4743E-02	3.3600E-03	2.8320E+00	7.1078E-01	2.5073E-02	1.3139E+01	6.2380E-02	1.9477E-01	6.1872E-01	9.0938E-01
1.7632E+00	8.0800E-03	2.2649E-01	7.2838E-02	6.5622E-02	3.6300E-03	2.4874E+00	8.9812E-01	2.7568E-02	1.2846E+01	5.9080E-02	2.4896E-01	6.9057E-01	4.7592E-01
6.1522E-01	1.6770E-02	4.0479E-01	5.9574E-02	2.4190E-02	5.7100E-03	1.4014E+00	1.3619E+00	2.2407E-02	8.8024E+00	5.9760E-02	1.2068E-03	3.8204E-01	7.9925E-01
2.2860E+00	1.3240E-02	3.5282E-01	5.8561E-02	2.5088E-02	4.6900E-03	2.6587E+00	1.0597E+00	3.3988E-03	8.7479E+00	5.9220E-02	2.4351E-01	1.6386E-01	9.8508E-01
5.8120E-01	1.5260E-02	3.3358E-01	6.0146E-02	3.2899E-02	4.2000E-03	3.3718E-01	6.5980E-01	8.1416E-03	8.1324E+00	6.4890E-02	4.9420E-02	6.4419E-01	6.1460E-01
2.1140E+00	1.2330E-02	3.0022E-01	6.1495E-02	5.7438E-02	4.2400E-03	8.7393E-01	9.9824E-01	1.0000E-02	1.0324E+01	6.6660E-02	5.7497E-02	1.6717E-01	2.4530E-01
6.0626E-01	1.0990E-02	2.5824E-01	5.9085E-02	3.2873E-02	4.2400E-03	7.6243E-01	2.4860E-01	1.3890E-02	1.0357E+01	5.8250E-02	1.8860E-02	6.1290E-01	4.2092E-01

8.2931E-01	1.1000E-02	2.6777E-01	6.0333E-02	6.5949E-02	4.1400E-03	2.6198E+00	1.5501E-01	2.8430E-02	1.0379E+01	7.1570E-02	1.1894E-01	3.3731E-01	4.5879E-01
2.0960E+00	2.2950E-02	3.2420E-01	3.2574E-02	3.5184E-02	4.1700E-03	1.9340E+00	8.3552E-01	1.6450E-02	8.8734E+00	5.6180E-02	6.5621E-02	2.6011E-01	8.4941E-01
3.0987E-01	1.9840E-02	3.1510E-01	3.3916E-02	7.9885E-04	4.4500E-03	1.8404E+00	5.7645E-01	6.8732E-03	8.4091E+00	5.0780E-02	1.7145E-01	1.4741E-01	7.7286E-01
1.9558E+00	1.9790E-02	3.1235E-01	3.6111E-02	1.6177E-02	3.6700E-03	1.8010E+00	5.8676E-01	1.8480E-02	8.0388E+00	4.8630E-02	3.3723E-02	3.8517E-01	4.9595E-01
1.2202E+00	1.6840E-02	2.9412E-01	2.2693E-02	6.0741E-02	2.5900E-03	1.1075E+00	5.9874E-01	2.8614E-02	1.0665E+01	5.1540E-02	2.0033E-01	1.7421E-02	7.1525E-01
2.1367E+00	1.2430E-02	2.8011E-01	2.4689E-02	1.6702E-02	2.5500E-03	2.8384E-01	3.5302E-01	2.8260E-02	9.6226E+00	4.9530E-02	2.3627E-01	6.8594E-01	7.1969E-01
1.7529E+00	1.6130E-02	3.2627E-01	2.3851E-02	6.1494E-03	1.8200E-03	2.8284E+00	9.1135E-01	1.6751E-02	1.0558E+01	5.5250E-02	1.0188E-01	1.8601E-01	6.9282E-01
1.8587E+00	1.6950E-02	2.5786E-01	1.1008E-01	1.2152E-02	3.6100E-03	4.3625E-01	6.6155E-01	1.6565E-02	1.5737E+01	5.2700E-02	1.9855E-01	1.8245E-02	9.2522E-01
1.7530E+00	1.5150E-02	5.2576E-01	8.3242E-02	1.7137E-02	5.2700E-03	1.0730E+00	5.2232E-03	1.0044E-02	1.3579E+01	6.6660E-02	1.1294E-01	2.3966E-01	9.9197E-01
5.9327E-01	1.2030E-02	3.7725E-01	7.9466E-02	6.0666E-02	4.1300E-03	1.9590E+00	9.6595E-01	2.1362E-03	1.1400E+01	4.0550E-02	3.1383E-02	6.4541E-01	8.8729E-01
1.8951E+00	1.2440E-02	5.5853E-01	8.2438E-02	4.0425E-02	7.4700E-03	1.3935E+00	3.1051E-01	2.8990E-02	1.2692E+01	7.1440E-02	1.3391E-01	5.9266E-01	5.7604E-01
2.2490E+00	4.0750E-02	3.6928E-01	4.9444E-02	6.4844E-02	6.2600E-03	1.6180E+00	6.9939E-01	6.8818E-03	1.2820E+01	8.5690E-02	1.9222E-02	5.3319E-01	9.0054E-01

2015						
Descriptor	**Product**	**Panelist**	**Rep**	**Product: panelist**	**Product:rep**	**Panelist:rep**
sour	9.08E-09	2.57E-25	8.62E-01	2.13E-01	4.33E-01	4.19E-03
floral	1.54E-06	4.81E-21	5.04E-02	1.62E-02	1.26E-01	9.01E-01
sweet	9.77E-06	4.66E-24	1.67E-02	1.44E-02	9.11E-01	2.39E-01
mouthfeel	9.78E-06	3.12E-13	2.57E-01	1.17E-02	2.53E-01	1.31E-01
apple.peach	1.28E-05	4.95E-10	3.64E-01	1.82E-01	8.08E-01	8.16E-01
body	5.21E-04	8.91E-07	7.35E-01	3.45E-03	1.88E-01	1.23E-05
honey	1.50E-03	1.09E-13	6.91E-01	1.52E-01	1.11E-01	6.87E-02
fruity_taste	2.53E-03	2.19E-14	8.14E-01	3.14E-05	3.12E-02	4.52E-02
exotic	7.32E-03	6.39E-15	3.34E-01	8.88E-04	5.84E-01	1.93E-01
lemon	1.23E-02	2.01E-17	2.21E-01	1.34E-01	5.05E-01	6.38E-01
smoky	1.24E-01	1.71E-14	4.63E-01	1.26E-02	9.07E-01	9.04E-01
bitter	2.83E-01	1.67E-14	1.20E-01	8.19E-02	4.26E-01	9.92E-01
oxidized_apple	7.75E-01	4.55E-20	2.62E-01	1.16E-02	1.97E-01	3.97E-01
green	9.12E-01	7.82E-19	6.18E-02	2.60E-02	9.33E-02	3.36E-01

2016						
Descriptor	**Product**	**Panelist**	**Rep**	**Product: panelist**	**Product:rep**	**Panelist:rep**
sweet	2.77E-46	5.67E-118	6.16E-01	7.97E-02	9.17E-01	4.14E-09
sour	5.71E-46	1.13E-89	2.86E-02	7.53E-07	7.38E-01	3.16E-03
mouthfeel	1.97E-33	6.27E-29	1.86E-02	3.24E-02	7.52E-01	4.88E-02
fruity_taste	1.22E-14	1.15E-93	9.55E-01	1.02E-01	7.37E-01	4.58E-05
oxidized_apple	2.33E-14	2.60E-91	6.85E-04	5.82E-03	8.93E-01	6.78E-01
apple.peach	3.67E-13	1.04E-60	1.80E-01	1.28E-02	4.89E-01	3.82E-03
smoky	4.07E-13	1.71E-115	3.52E-02	1.74E-04	1.66E-01	4.96E-01
body	1.56E-12	1.87E-83	5.69E-01	3.51E-05	5.28E-01	1.22E-01
lemon	2.16E-11	1.32E-66	2.87E-02	1.03E-02	3.74E-01	1.31E-03
exotic	3.20E-11	6.61E-62	2.24E-02	1.06E-10	5.40E-02	3.34E-04
floral	2.81E-09	2.90E-66	6.33E-01	3.02E-04	3.10E-01	4.36E-02
honey	1.48E-05	1.19E-90	7.87E-05	9.30E-02	5.50E-01	8.44E-01
bitter	4.90E-05	1.91E-85	1.73E-01	6.21E-06	5.47E-01	1.11E-04
green	1.06E-01	3.94E-111	5.44E-01	5.61E-02	9.04E-01	9.42E-02

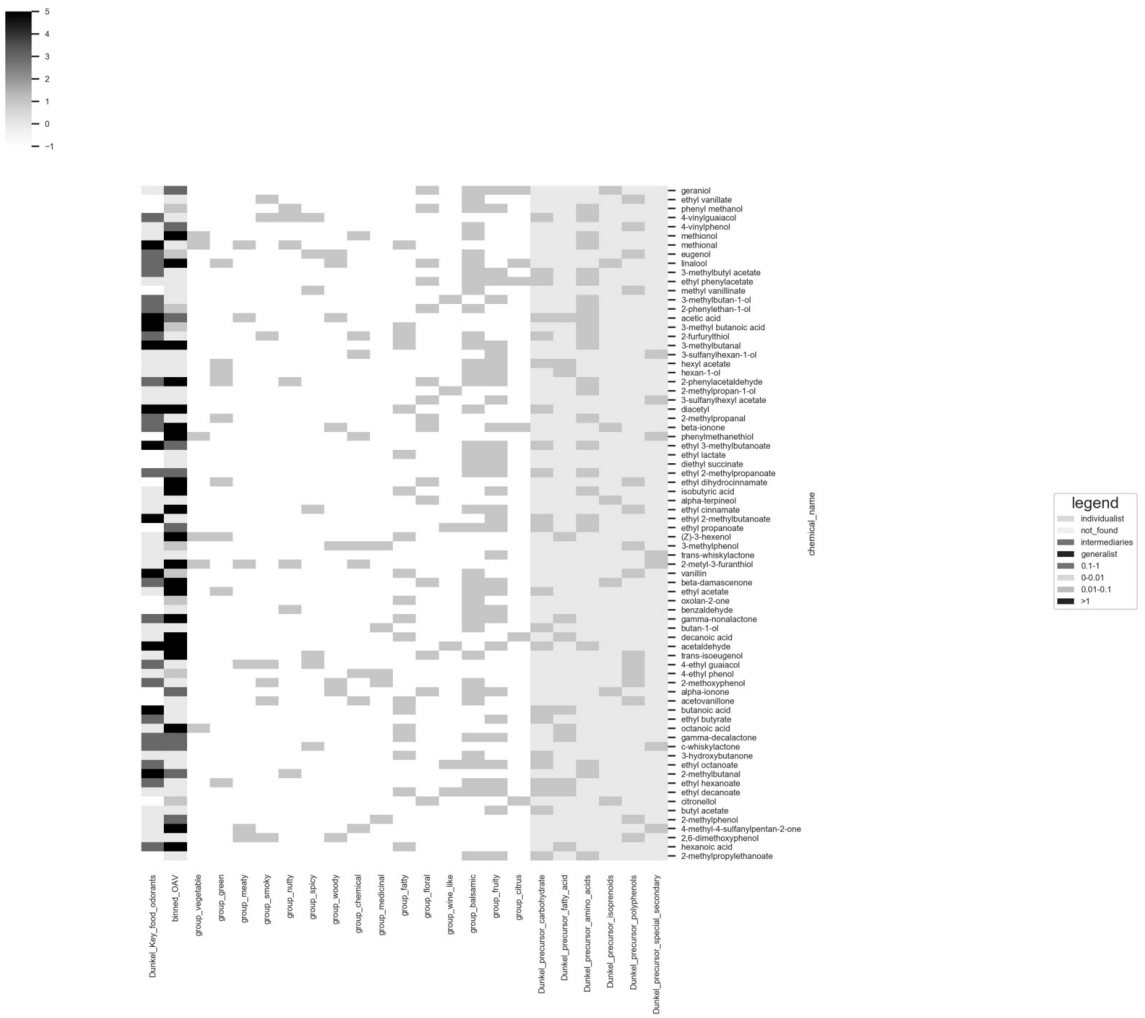


Figure S6 Compound origin shown as precursors (Dunkel et al., 2014) in a presence-absence matrix with odour relationships.

Supplementary tables

2016				
Axis	**Df**	**ChiSquare**	**F**	**Pr(>F)**
CCA1	1	1.28E-02	8.54E+01	1.00E-03
CCA2	1	6.02E-03	4.01E+01	1.00E-03
CCA3	1	3.47E-03	2.31E+01	1.00E-03
CCA4	1	2.47E-03	1.64E+01	7.40E-02
CCA5	1	1.68E-03	1.12E+01	9.69E-01
CCA6	1	1.41E-03	9.38E+00	1.00E+00
CCA7	1	1.07E-03	7.09E+00	1.00E+00
CCA8	1	9.22E-04	6.13E+00	1.00E+00
CCA9	1	8.80E-04	5.85E+00	1.00E+00
CCA10	1	6.98E-04	4.64E+00	1.00E+00
CCA11	1	3.46E-04	2.30E+00	1.00E+00
CCA12	1	2.65E-04	1.77E+00	1.00E+00
CCA13	1	2.45E-04	1.63E+00	1.00E+00
2015				
Axis	**Df**	**ChiSquare**	**F**	**Pr(>F)**
CCA1	1	1.13E-02	1.65E+01	1.00E-03
CCA2	1	6.88E-03	1.01E+01	1.00E-03
CCA3	1	4.39E-03	6.43E+00	2.80E-02
CCA4	1	2.07E-03	3.03E+00	9.58E-01
CCA5	1	1.66E-03	2.43E+00	9.99E-01
CCA6	1	1.23E-03	1.81E+00	1.00E+00
CCA7	1	8.67E-04	1.27E+00	1.00E+00
CCA8	1	6.58E-04	9.65E-01	1.00E+00
CCA9	1	4.18E-04	6.13E-01	1.00E+00
CCA10	1	3.66E-04	5.37E-01	1.00E+00
CCA11	1	1.40E-04	2.05E-01	1.00E+00
CCA12	1	8.30E-05	1.21E-01	1.00E+00
CCA13	1	2.50E-05	3.61E-02	1.00E+00

Table S1. Description of the sensory attributes and their reference standard composition.

Attribute	**Description (Reference-Standard composition for 100 mL)**
citrus	fresh lemon juice (10 mL)
exotic fruit	passionfruit juice (1.5mL, Monin)
yellow fruit	peach juice (10 mL , Granini) + apple juice (6 mL, ReWe)
oxidized apple	apple juice (15 mL, ReWe) + acetaldehyde (150 mg/L)
green	trans-2-hexenal (35 µL, 10 mg / 10 ml) + gooseberry pickle (5 mL, Hainich) + green apple (3 g, Granny Smith)
floral	linalool (2 drops , 1 mg/ mL) + geraniol (1 drop, 1 mg/1 mL)
honey	flower honey (2 g, Langnese)
smoky	single malt whiskey (6 drops, Laphroaig 10 years) + 4-vinylguaiacol (150 µg/L)

Table S2. Details of the chemical compounds metadata.

Table S3. Details of the chemical compounds. Measured chemical compounds (mg/L, if not measured or under threshold then filled with random value of max the threshold).

Table S4. Sensory panel performance (p-values) for both years.

Table S5. Constrained correspondence analysis (CCA) permutation test for the axis significance for 2016 and 2015.

Table S6. Scaled chemical compound raw values pairwise correlation edges with rho $\in$]-0.15,0.15[

Node 1	Node 2	Left-tail rho	Right-tail rho	Median rho
2-methylbutanal	3-sulfanylhexyl acetate	-1,595E-01	-6,003E-01	-3,873E-01
linalool	phenyl methanol	5,727E-01	1,542E-01	3,693E-01
4-vinylguaiacol	linalool	5,806E-01	1,691E-01	3,837E-01
4-vinylphenol	eugenol	6,314E-01	2,447E-01	4,473E-01
methional	phenyl methanol	6,444E-01	2,508E-01	4,613E-01
phenyl methanol	methionol	6,846E-01	2,984E-01	5,005E-01
4-vinylphenol	3-methylbutyl acetate	6,87E-01	3,199E-01	5,157E-01
4-vinylphenol	ethyl phenylacetate	6,921E-01	3,257E-01	5,236E-01
4-vinylphenol	linalool	7,009E-01	3,37E-01	5,271E-01
ethyl 3-methylbutanoate	ethyl hexanoate	7,292E-01	3,775E-01	5,646E-01
4-methyl-4-sulfanylpentan-2-one	hexanoic acid	7,326E-01	3,938E-01	5,723E-01
hexan-1-ol	hexyl acetate	7,364E-01	3,969E-01	5,755E-01
4-methyl-4-sulfanylpentan-2-one	2,6-dimethoxyphenol	7,324E-01	3,941E-01	5,758E-01
4-methyl-4-sulfanylpentan-2-one	2-methylpropylethanoate	7,53E-01	4,233E-01	5,994E-01
2-phenylacetaldehyde	hexyl acetate	7,545E-01	4,247E-01	6,025E-01
ethyl decanoate	ethyl hexanoate	7,581E-01	4,343E-01	6,086E-01
citronellol	ethyl hexanoate	7,642E-01	4,400E-01	6,14E-01
ethyl vanillate	4-vinylphenol	7,835E-01	4,845E-01	6,471E-01
butan-1-ol	acetaldehyde	7,81E-01	5,004E-01	6,486E-01
ethyl vanillate	phenyl methanol	7,900E-01	4,918E-01	6,538E-01
4-vinylphenol	phenyl methanol	7,987E-01	5,124E-01	6,665E-01

ethyl 2-methylbutanoate	ethyl propanoate	8,079E-01	5,249E-01	6,769E-01
4-ethyl guaiacol	butan-1-ol	8,036E-01	5,428E-01	6,84E-01
trans-isoeugenol	butan-1-ol	8,054E-01	5,41E-01	6,855E-01
4-vinylguaiacol	phenyl methanol	8,164E-01	5,467E-01	6,923E-01
4-ethyl phenol	butan-1-ol	8,135E-01	5,546E-01	6,943E-01
ethyl vanillate	4-vinylguaiacol	8,174E-01	5,491E-01	6,968E-01
4-ethyl phenol	acetaldehyde	8,175E-01	5,729E-01	7,032E-01
ethyl 2-methylpropanoate	ethyl lactate	8,289E-01	5,731E-01	7,14E-01
decanoic acid	acetaldehyde	8,319E-01	5,814E-01	7,144E-01
4-ethyl guaiacol	acetaldehyde	8,359E-01	5,966E-01	7,274E-01
ethyl dihydrocinnamate	alpha-terpineol	8,381E-01	6,034E-01	7,305E-01
ethyl 2-methylbutanoate	(Z)-3-hexenol	8,421E-01	6,01E-01	7,33E-01
trans-isoeugenol	acetaldehyde	8,455E-01	6,088E-01	7,366E-01
(Z)-3-hexenol	ethyl propanoate	8,433E-01	6,063E-01	7,383E-01
beta-damascenone	oxolan-2-one	8,466E-01	6,094E-01	7,415E-01
phenylmethanethiol	beta-ionone	8,497E-01	6,16E-01	7,439E-01
4-ethyl guaiacol	decanoic acid	8,475E-01	6,29E-01	7,474E-01
2-furfurylthiol	3-methyl butanoic acid	8,517E-01	6,216E-01	7,491E-01
ethyl cinnamate	ethyl dihydrocinnamate	8,522E-01	6,273E-01	7,511E-01
eugenol	3-methylbutyl acetate	8,559E-01	6,373E-01	7,571E-01
beta-damascenone	ethyl acetate	8,589E-01	6,36E-01	7,594E-01
3-methylbutanal	3-methyl butanoic acid	8,611E-01	6,396E-01	7,626E-01
diethyl succinate	ethyl lactate	8,649E-01	6,474E-01	7,656E-01
ethyl dihydrocinnamate	isobutyric acid	8,633E-01	6,486E-01	7,679E-01
trans-isoeugenol	decanoic acid	8,634E-01	6,537E-01	7,681E-01
eugenol	linalool	8,688E-01	6,581E-01	7,727E-01
4-ethyl phenol	decanoic acid	8,717E-01	6,67E-01	7,797E-01
3-methylbutanal	2-furfurylthiol	8,764E-01	6,752E-01	7,861E-01
c-whiskylactone	3-hydroxybutanone	8,755E-01	6,748E-01	7,873E-01

linalool	3-methylbutyl acetate	8,777E-01	6,785E-01	7,88E-01
eugenol	ethyl phenylacetate	8,753E-01	6,782E-01	7,883E-01
c-whiskylactone	ethyl octanoate	8,81E-01	6,84E-01	7,916E-01
trans-isoeugenol	4-ethyl guaiacol	8,822E-01	6,898E-01	7,962E-01
4-vinylphenol	4-vinylguaiacol	8,827E-01	6,903E-01	7,965E-01
ethyl octanoate	3-hydroxybutanone	8,824E-01	6,925E-01	7,986E-01
decanoic acid	butan-1-ol	8,853E-01	6,958E-01	7,99E-01
oxolan-2-one	ethyl acetate	8,834E-01	6,947E-01	8,001E-01
2-phenylacetaldehyde	hexan-1-ol	8,859E-01	6,981E-01	8,032E-01
trans-isoeugenol	4-ethyl phenol	8,883E-01	7,048E-01	8,059E-01
acetovanillone	alpha-ionone	8,894E-01	7,057E-01	8,072E-01
4-ethyl phenol	4-ethyl guaiacol	8,919E-01	7,124E-01	8,114E-01
2,6-dimethoxyphenol	hexanoic acid	8,967E-01	7,227E-01	8,185E-01
alpha-terpineol	isobutyric acid	8,977E-01	7,262E-01	8,215E-01
linalool	ethyl phenylacetate	8,987E-01	7,292E-01	8,231E-01
2,6-dimethoxyphenol	2-methylpropylethanoate	9,022E-01	7,344E-01	8,263E-01
2-methylphenol	butyl acetate	9,029E-01	7,38E-01	8,295E-01
octanoic acid	ethyl butyrate	9,038E-01	7,406E-01	8,313E-01
2-phenylethan-1-ol	3-methylbutan-1-ol	9,064E-01	7,467E-01	8,352E-01
ethyl cinnamate	isobutyric acid	9,062E-01	7,47E-01	8,367E-01
ethyl 2-methylpropanoate	diethyl succinate	9,06E-01	7,487E-01	8,371E-01
trans-whiskylactone	3-methylphenol	9,098E-01	7,535E-01	8,399E-01
methional	methionol	9,124E-01	7,596E-01	8,43E-01
citronellol	ethyl decanoate	9,134E-01	7,637E-01	8,476E-01
2-methylpropylethanoate	hexanoic acid	9,194E-01	7,783E-01	8,564E-01
ethyl cinnamate	alpha-terpineol	9,219E-01	7,841E-01	8,604E-01
ethyl phenylacetate	3-methylbutyl acetate	9,212E-01	7,842E-01	8,615E-01

Table S7. OAV scaled chemical compound pairwise correlation edges with rho ∈]-0.15,0.15[

Node 1	Node 2	Left-tail rho	Right-tail rho	Median rho
3-sulfanylhexan-1-ol	beta-damascenone	-3,451E-01	-7,006E-01	-5,31E-01
methional	3-hydroxybutanone	-3,479E-01	-6,813E-01	-5,236E-01
2-phenylethan-1-ol	3-hydroxybutanone	-3,314E-01	-6,927E-01	-5,194E-01
methionol	3-hydroxybutanone	-3,259E-01	-6,679E-01	-5,039E-01
2-phenylacetaldehyde	3-hydroxybutanone	-3,046E-01	-6,67E-01	-4,944E-01
2-methylbutanal	3-sulfanylhexyl acetate	-2,833E-01	-6,778E-01	-4,866E-01
ethyl lactate	hexyl acetate	-2,759E-01	-6,71E-01	-4,864E-01
2-phenylacetaldehyde	beta-damascenone	-3,052E-01	-6,518E-01	-4,84E-01
ethyl 2-methylpropanoate	3-methylbutyl acetate	-2,800E-01	-6,501E-01	-4,699E-01
2-methylbutanal	3-sulfanylhexan-1-ol	-2,817E-01	-6,311E-01	-4,579E-01
methionol	ethyl octanoate	-2,622E-01	-6,393E-01	-4,541E-01
2-methylpropylethanoate	3-methyl butanoic acid	-2,685E-01	-6,18E-01	-4,439E-01
2-phenylacetaldehyde	2-methyl-3-furanthiol	-2,59E-01	-6,224E-01	-4,438E-01
methionol	ethyl butyrate	-2,608E-01	-6,196E-01	-4,423E-01
methional	ethyl octanoate	-2,532E-01	-6,036E-01	-4,336E-01
2-phenylacetaldehyde	diethyl succinate	-2,483E-01	-6,132E-01	-4,326E-01
methional	beta-damascenone	-2,194E-01	-6,222E-01	-4,32E-01
methional	2-methylbutanal	-2,486E-01	-6,061E-01	-4,29E-01
methionol	butan-1-ol	-2,473E-01	-5,98E-01	-4,262E-01
2-methylbutanal	4-vinylguaiacol	-2,38E-01	-5,923E-01	-4,185E-01
isobutyric acid	hexyl acetate	-2,201E-01	-6,131E-01	-4,183E-01
ethyl phenylacetate	ethyl 2-methylpropanoate	-2,358E-01	-5,921E-01	-4,158E-01
2-phenylacetaldehyde	2-methylbutanal	-2,354E-01	-5,766E-01	-4,063E-01

2-phenylacetaldehyde	ethyl butyrate	-2,227E-01	-5,919E-01	-4,063E-01
methional	ethyl butyrate	-2,138E-01	-5,869E-01	-4,058E-01
methional	butan-1-ol	-2,25E-01	-5,777E-01	-4,038E-01
2-methylpropylethanoate	2-phenylethan-1-ol	-2,22E-01	-5,692E-01	-3,983E-01
3-sulfanylhexan-1-ol	2-methyl-3-furanthiol	-1,978E-01	-5,906E-01	-3,963E-01
2-phenylethan-1-ol	butan-1-ol	-2,205E-01	-5,683E-01	-3,961E-01
methional	4-methyl-4-sulfanylpentan-2-one	-1,752E-01	-5,905E-01	-3,92E-01
2-methylpropylethanoate	methionol	-1,915E-01	-5,704E-01	-3,861E-01
2-phenylethan-1-ol	ethyl butyrate	-2,075E-01	-5,571E-01	-3,852E-01
2-phenylacetaldehyde	ethyl lactate	-1,691E-01	-5,846E-01	-3,831E-01
2-methylpropan-1-ol	3-hydroxybutanone	-1,992E-01	-5,61E-01	-3,822E-01
4-methyl-4-sulfanylpentan-2-one	4-vinylguaiacol	-1,623E-01	-5,85E-01	-3,795E-01
3-methyl butanoic acid	3-hydroxybutanone	-1,818E-01	-5,679E-01	-3,775E-01
2-methylpropylethanoate	ethyl 2-methylbutanoate	-1,878E-01	-5,608E-01	-3,774E-01
3-sulfanylhexan-1-ol	ethyl acetate	-1,734E-01	-5,528E-01	-3,664E-01
ethyl 2-methylbutanoate	3-methylbutyl acetate	-1,606E-01	-5,578E-01	-3,655E-01
beta-damascenone	methionol	-1,559E-01	-5,651E-01	-3,653E-01
4-vinylguaiacol	ethyl 2-methylpropanoate	-1,558E-01	-5,471E-01	-3,571E-01
methional	2-methylpropylethanoate	-1,600E-01	-5,454E-01	-3,554E-01
3-sulfanylhexan-1-ol	ethyl octanoate	-1,574E-01	-5,356E-01	-3,488E-01
2-methylbutanal	2-methylpropan-1-ol	-1,783E-01	-5,152E-01	-3,475E-01
2-methylpropan-1-ol	ethyl octanoate	-1,703E-01	-5,169E-01	-3,438E-01
ethyl 2-methylpropanoate	octanoic acid	-1,695E-01	-5,174E-01	-3,379E-01
ethyl phenylacetate	oxolan-2-one	-1,595E-01	-5,217E-01	-3,376E-01
3-sulfanylhexan-1-ol	oxolan-2-one	-1,503E-01	-5,193E-01	-3,322E-01
butan-1-ol	ethyl octanoate	5,102E-01	1,651E-01	3,357E-01

2-phenylacetaldehyde	2-methylpropan-1-ol	5,113E-01	1,719E-01	3,389E-01
oxolan-2-one	diethyl succinate	5,184E-01	1,75E-01	3,47E-01
2-phenylacetaldehyde	2-phenylethan-1-ol	5,296E-01	1,582E-01	3,476E-01
3-sulfanylhexan-1-ol	4-vinylguaiacol	5,397E-01	1,541E-01	3,487E-01
2-methylbutanal	oxolan-2-one	5,162E-01	1,818E-01	3,491E-01
methional	4-vinylguaiacol	5,422E-01	1,559E-01	3,524E-01
2-methylpropylethanoate	ethyl butyrate	5,322E-01	1,69E-01	3,527E-01
ethyl 2-methylpropanoate	oxolan-2-one	5,329E-01	1,74E-01	3,542E-01
butanoic acid	3-methylbutyl acetate	5,36E-01	1,76E-01	3,543E-01
2-phenylacetaldehyde	methionol	5,404E-01	1,584E-01	3,564E-01
2-methylbutanal	3-hydroxybutanone	5,493E-01	1,601E-01	3,573E-01
2-methylpropylethanoate	hexanoic acid	5,35E-01	1,92E-01	3,613E-01
2-methylpropylethanoate	3-hydroxybutanone	5,652E-01	1,605E-01	3,639E-01
2-phenylacetaldehyde	3-sulfanylhexan-1-ol	5,429E-01	1,758E-01	3,668E-01
2-methylbutanal	ethyl octanoate	5,422E-01	1,84E-01	3,669E-01
ethyl lactate	ethyl butyrate	5,677E-01	1,547E-01	3,674E-01
methional	phenyl methanol	5,796E-01	1,518E-01	3,733E-01
methional	linalool	5,545E-01	1,815E-01	3,735E-01
vanillin	beta-damascenone	5,791E-01	1,501E-01	3,743E-01
methionol	2-methylpropan-1-ol	5,622E-01	1,895E-01	3,755E-01
octanoic acid	3-methylbutyl acetate	5,545E-01	1,969E-01	3,773E-01
beta-damascenone	oxolan-2-one	5,675E-01	1,918E-01	3,796E-01
ethyl 2-methylbutanoate	3-methyl butanoic acid	5,793E-01	1,817E-01	3,847E-01
4-vinylguaiacol	linalool	5,707E-01	1,963E-01	3,865E-01
2-methylbutanal	2-methyl-3-furanthiol	5,697E-01	2,002E-01	3,879E-01
ethyl lactate	3-hydroxybutanone	5,969E-01	1,616E-01	3,885E-01
methional	2-methylpropan-1-ol	5,61E-01	2,049E-01	3,894E-01

4-vinylguaiacol	eugenol	5,700E-01	1,993E-01	3,903E-01
geraniol	linalool	5,857E-01	1,941E-01	3,907E-01
trans-isoeugenol	eugenol	5,983E-01	1,826E-01	3,945E-01
2-methyl-3-furanthiol	oxolan-2-one	5,755E-01	2,151E-01	3,949E-01
isobutyric acid	ethyl lactate	6,059E-01	1,775E-01	3,997E-01
4-methyl-4-sulfanylpentan-2-one	3-hydroxybutanone	6,02E-01	1,848E-01	4,026E-01
acetovanillone	ethyl lactate	5,79E-01	2,203E-01	4,049E-01
3-sulfanylhexan-1-ol	methionol	5,98E-01	2,028E-01	4,056E-01
2-methylbutanal	diethyl succinate	5,911E-01	2,25E-01	4,087E-01
hexanoic acid	3-methylbutyl acetate	5,938E-01	2,304E-01	4,112E-01
ethyl 3-methylbutanoate	ethyl 2-methylbutanoate	5,968E-01	2,288E-01	4,162E-01
4-vinylphenol	ethyl phenylacetate	6,144E-01	2,111E-01	4,183E-01
ethyl 2-methylbutanoate	ethyl 2-methylpropanoate	6,02E-01	2,321E-01	4,21E-01
ethyl octanoate	3-hydroxybutanone	5,99E-01	2,318E-01	4,225E-01
2-methyl-3-furanthiol	ethyl 2-methylpropanoate	6,223E-01	2,165E-01	4,235E-01
4-vinylguaiacol	ethyl phenylacetate	5,989E-01	2,448E-01	4,268E-01
2-phenylacetaldehyde	methional	6,046E-01	2,347E-01	4,27E-01
2-methylpropylethanoate	3-methylbutyl acetate	6,241E-01	2,186E-01	4,294E-01
ethyl 2-methylpropanoate	diethyl succinate	6,204E-01	2,316E-01	4,299E-01
4-vinylphenol	3-methylbutyl acetate	6,167E-01	2,401E-01	4,331E-01
eugenol	phenyl methanol	6,286E-01	2,247E-01	4,345E-01
methional	3-sulfanylhexan-1-ol	6,185E-01	2,459E-01	4,389E-01
ethyl 3-methylbutanoate	ethyl 2-methylpropanoate	6,231E-01	2,471E-01	4,42E-01
ethyl vanillate	ethyl lactate	6,216E-01	2,39E-01	4,421E-01
4-vinylphenol	4-vinylguaiacol	6,383E-01	2,299E-01	4,438E-01
3-methyl butanoic acid	methionol	6,146E-01	2,622E-01	4,448E-01
ethyl vanillate	phenyl methanol	6,394E-01	2,325E-01	4,493E-01
ethyl butyrate	3-hydroxybutanone	6,339E-01	2,766E-01	4,646E-01

2-methylbutanal	beta-damascenone	6,495E-01	2,808E-01	4,708E-01
3-methyl butanoic acid	2-phenylethan-1-ol	6,53E-01	3,063E-01	4,821E-01
beta-damascenone	ethyl acetate	6,715E-01	2,97E-01	4,903E-01
methional	3-methyl butanoic acid	6,679E-01	3,043E-01	4,947E-01
beta-damascenone	diethyl succinate	6,717E-01	3,09E-01	4,967E-01
butan-1-ol	3-hydroxybutanone	6,824E-01	3,167E-01	5,037E-01
methional	2-phenylethan-1-ol	6,622E-01	3,466E-01	5,077E-01
2-phenylethan-1-ol	methionol	6,803E-01	3,43E-01	5,172E-01
2-methyl-3-furanthiol	diethyl succinate	6,844E-01	3,375E-01	5,178E-01
acetovanillone	ethyl vanillate	7,414E-01	4,000E-01	5,794E-01
2-methyl-3-furanthiol	beta-damascenone	7,377E-01	4,031E-01	5,822E-01
methional	methionol	7,537E-01	4,336E-01	6,027E-01
ethyl phenylacetate	3-methylbutyl acetate	7,709E-01	4,557E-01	6,263E-01

Conclusion

In this thesis the native wine microbial communities were studied to reveal their effects on chemical and sensory composition of Riesling wine. The microbial community profiles, effects and interactions from both vineyard and winery environments were investigated by applying high-throughput technologies. This thesis demonstrated that both vineyard and winery microbial communities play immensely significant roles in wine. The microbial communities within the fermenting must were strongly influenced by vineyard of origin. Additionally, winery microbial communities were found to play a prominent role as well, thus the interaction between winery and vineyard communities consequently translated into diverse sensory properties.

Previous studies linking the effect of native microbial communities to sensory relevant aroma compounds with their interactive properties have been vastly unsuccessful to date. As wine is a product of complex microbial communities, studies that base their approach on relatively few isolated strains or chemical compounds are not sufficient for fully understanding this complex outcome. This thesis has demonstrated that with combining modern untargeted high-throughput technologies and statistical approaches it is possible to look into samples *in situ* in the actual natural environment. This methodology allowed finding novel microbial and chemical patterns which could be further tested with targeted studies. Additionally, this thesis showed that the statistical approaches also allow to "fine-tune" older technologies to produce higher turnover.

The major achievement in this thesis was the realisation how multiple simultaneous research directions are possible by combining untargeted approaches. In addition to deconstructing the microbial community composition in complex natural environment, the shotgun metagenomic data was leveraged on, to undertake the first functional analysis of the microbial community during wine fermentation. In the future, multi-omics approaches will be essential for fully discovering the complexity of biological networks, where microbes, host and chemical compounds interact with to human sensory perceptions. These here developed approaches will also benefit any such industry that works with complex biological interactions. The next step is to expand these early results to multi-omics studies with larger sample sets, and integrate other 'omics tools in combination of profiling the whole genomic content using true metagenomic approaches. This approach could potentially provide information about the biological processes such as metabolic pathways. Such work in wine is currently underway and will be most informative with regards to the groundwork approaches and methods that have been developed during this thesis.

Publications related to thesis

Sirén K., Fischer U., Vestner J. (2019). Automated supervised learning pipeline for non-targeted GC-MS data analysis. Analytical Chimica Acta: X 1, 100005.

Sirén, K., Mak, S. S. T., Fischer, U., Hansen, L. H., and Gilbert, M. T. P. (2019). Multi-omics and potential applications in wine production. Current Opinion in Biotechnology 56, 172–178.

(Under review / in preparation)

Sirén* K., Mak* S.S.T., Melkonian C., Carøe C., Swiegers J.H., Molenaar D., Fischer U., Gilbert M.T.P. (2019) *in review*. Taxonomic and functional characterisation of the microbial community during spontaneous in vitro fermentation of Riesling must. *in review*

Sirén K., Oliveira I., Ferreira F., Fischer U. (2019) *in prep*. Impact of vineyard and cellar microbiota on the distinction of different Riesling vineyards based on aroma and sensory analysis. *in prep*

Mak* S.S.T, Sirén* K., Carøe C., Gopalakrishnan S., Ellegaard M.R., Klincke F., Hansen J. H, Fischer U., Gilbert M.T.P. (2019) *in prep* Comparison of DNA extraction methods for use in fungal diversity analyses of Riesling during alcoholic fermentation. *in prep*

Scientific Posters/Talks related to thesis

Using novel supervised learning to find key features from non-targeted GC-MS data. Chemometrics in Analytical Chemistry Halifax June 2018. Talk.

Impact of vineyard versus cellar microbiota on the distinction of different Riesling vineyards based on aroma and sensory analysis, Macrowine Zaragoza, May 2018. Talk.

Spontaneous fermentation dynamics of indigenous yeast populations and their effect on the sensory properties of Riesling. XI International Terroir Congress in Oregon, July 2016. Talk.

Spontaneous fermentation dynamics of indigenous yeast populations and their effect on the sensory properties of Riesling. Macrowine in Changings, June 2016. Poster.

Enhancing varietal aroma in cool-climate varietal wines: impact of Pichia and Saccharomyces yeasts. ICCWS in Brighton, May 2016. Poster.

Curriculum Vitae

Kimmo Sirén

Education

Marie Curie PhD student / Microwine — DLR Rheinpfalz, Institute for Viticulture & Oenology, Germany & TU Kaiserslautern, Kaiserslautern, Germany from 9/2015

MSc. / Enology & Viticulture — FH Geisenheim University, Germany & ISVV Bordeaux &Montpellier SupAgro, France, Full Scholarship, Vinifera EuroMaster, 2014

BSc. / Mathematics — University of Helsinki, Finland, 2012

Level 4 Diploma — Wine & Spirit Education Trust (WSET), 2012

Level 3 Advanced — Wine & Spirit Education Trust (WSET), passed with Merit, 2009

Academic Work Experience

VISITING SCHOLAR, LEIBNIZ-LSB@TUM, GERMANY 10/2018

Metabolomics study on wine using MS-SWATH approach method established and multi-omics data integration

with Andreas Dunkel in Dr Thomas Hoffmann's Lab.

VISITING SCHOLAR, AUSTRALIAN WINE RESEARCH INSTITUTE, AUSTRALIA 2-3/2018

Metabolomics study on Riesling spontaneous fermentation kinetics using LC-MS/MS raw data and machine learning pipelines in Dr. Herderich's Lab.

VISITING SCHOLAR, DR FERREIRA'S LAB, SPAIN 10-11/2016

Targeted aroma study on Riesling wines using GC-MS based approaches and extracting information using supervised learning in Dr. Ferreira's Lab.

VISITING SCHOLAR, WASHINGTON STATE UNIVERSITY, USA 4-7/2014

Riesling research on phenolics, aroma compounds with chemical and sensory analysis in Dr. Jim Harbertson's Lab.

Workshops attended

EMBL-EBI Cambridge 30 October - 2 November 2017. Metabolomics Workflows

EMBL Heidelberg 23 - 24 May 2016. Next Generation Sequencing: Whole Genome Sequencing Library Preparation

Languages

Finnish, native; English fluent; German, excellent; Swedish good; French good;

Japanese basics; Python, R, Matlab, Bash: fluent

Scientific Awards

2nd prize for the best poster in ICCWS
http://www.iccws2016.com
1st prize for Jung-Sensoriker 2014, Deutsche Gesellschaft für Sensorik. DGSens Ev.
http://www.dgsens.de/foerdermassnahmen.html
Vinifera EuroMaster, Full Scholarship for 2012-2014
http://vinifera-euromaster.eu/